Sparse Image Reconstruction in Computed Tomography

Jørgensen, Jakob Sauer

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Sparse Image Reconstruction in Computed Tomography

Jakob Sauer Jørgensen

Kongens Lyngby 2013
PHD-2013-293

Technical University of Denmark
Applied Mathematics and Computer Science
Building 303B, 2800 Kongens Lyngby, Denmark
Phone +45 45253021
compute@compute.dtu.dk
www.compute.dtu.dk

PHD: ISSN 0909-3192

Summary

In recent years, increased focus on the potentially harmful effects of x-ray computed tomography (CT) scans, such as radiation-induced cancer, has motivated research on new low-dose imaging techniques. Sparse image reconstruction methods, as studied for instance in the field of compressed sensing (CS), have shown significant empirical potential for this purpose. For example, total variation regularized image reconstruction has been shown in some cases to allow reducing x-ray exposure by a factor of 10 or more, while maintaining or even improving image quality compared to conventional reconstruction methods.

However, the potential in CT has mainly been demonstrated in individual proof-of-concept studies, from which it is hard to distill general conditions for when sparse reconstruction methods perform well. As a result, there is a fundamental lack of understanding of the effectiveness and limitations of sparse reconstruction methods in CT, in particular in a quantitative sense. For example, relations between image properties such as contrast, structure and sparsity, tolerable noise levels, sufficient sampling levels, the choice of sparse reconstruction formulation and the achievable image quality remain unclear. This is a problem of high practical concern, because the large scale of CT problems makes detailed exploration of the parameter space very time-consuming. Due to the limited quantitative understanding, sparse reconstruction has not yet become the method of choice in practical CT applications.

This thesis takes a systematic approach toward establishing quantitative understanding of conditions for sparse reconstruction to work well in CT. A general framework for analyzing sparse reconstruction methods in CT is introduced and two sets of computational tools are proposed:

1. An optimization algorithm framework enabling easy derivation of algorithms for sparse reconstruction problems, and
2. Tools for characterizing sparse reconstruction in CT, i.e., establishing relations between parameters governing reconstruction quality.

The flexibility of the optimization algorithm framework is demonstrated by constructing convergent optimization algorithms for a range of sparse reconstruction problems of interest to CT. The practical usefulness of the framework is shown through case studies of the effectiveness of specific sparse reconstruction problems in tomographic reconstruction.

The characterization methods proposed in the thesis focus on the role of image sparsity for the level of sampling required for accurate CT reconstruction. While a relation between sparsity and sampling is motivated by CS, no theoretical guarantees of accurate sparse reconstruction are known for CT. In simulation studies, a sparsity-sampling relation is established in CT. This enables quantification of the undersampling allowed by sparse reconstruction methods.

Both the prototyping framework and the characterization methods add to the understanding of sparse reconstruction methods in CT and serve as initial contributions to a general set of computational characterization tools. Thus, the thesis contributions help advance sparse reconstruction methods toward routine use in practical applications of tomographic reconstruction, such as low-dose CT.

Resumé

Igennem de senere år har et øget fokus på potentielt skadelige effekter af CT (computed tomography) scanning, såsom strålingsinduceret cancer, motiveret forskning i nye lav-dosis billeddannelsesteknikker. Metoder baseret på algoritmer til sparse rekonstruktion, som f.eks. studeret inden for compressed sensing (CS), har vist betydeligt potentiale i denne anvendelse. For eksempel har total variation-regulariseret billedrekonstruktion vist sig i nogle tilfælde at kunne reducere røntgenstrålingseksponering med en faktor 10 eller mere, men stadig give samme eller bedre billedkvalitet som konventionelle rekonstruktionsmetoder.

Potentialet for at anvende disse teknikker inden for CT er imidlertid hovedsageligt blevet demonstreret i enkeltstående proof-of-concept studier, der ikke gør det klart hvilke generelle betingelser der skal være opfyldt, før sparse rekonstruktionsmetoder er velegnede. Derfor mangles fortsat en fundamental forståelse for effektiviteten og begrænsningerne af disse teknikkers anvendelse inden for CT, særligt i en kvantitativ forstand. Eksempelvis er det uklart hvilke sammenhænge der er mellem den opnåelige billedkvalitet, valg af sparse rekonstruktionsmetode, støjniveau, mængde af måledata og billedegenskaber såsom kontrast, struktur og sparsitet. Fra et praktisk synspunkt udgør dette et stort problem, da CT-rekonstruktion er meget beregningstungt og det derfor er særdeles tidskrævende at undersøge konsekvensen af parameter-valg. Som følge af den begrænsede kvantitative forståelse er sparse rekonstruktionsmetoder endnu ikke udbredt i praktiske anvendelser af CT.

Denne afhandling søger systematisk at etablere kvantitativ forståelse for de betingelser der afgør anvendeligheden af sparse rekonstruktionsmetoder i CT. Et generelt framework til analyse af sparse rekonstruktionsmetoder i CT introduceres og to typer beregningsmæssige analyseværktøjer foreslås:

1. Et framework af optimeringsalgoritmer til simpel udledning af algoritmer for sparse rekonstruktionsproblemer, og
2. Metoder til at karakterisere sammenhænge mellem forskellige problemparametre og disses indflydelse på kvaliteten af rekonstruerede billeder.

Fleksibiliteten af det foreslåede framework af optimeringsalgoritmer illustreres ved at konstruere konvergente optimeringsalgoritmer for en række sparse rekonstruktionsproblemer med relevans for anvendelsen i CT. Den praktiske anvendelighed af frameworket demonstreres gennem case studies, der undersøger effektiviteten af specifikke sparse rekonstruktionsproblemer i tomografisk rekonstruktion.

De foreslåede karakteriseringsmetoder fokuserer på sammenhængen mellem et billedes sparsitet og mængden af CT-måledata der kræves for at opnå en nøjagtig CT-rekonstruktion. En potentiel relation mellem billedsparsitet og mængden af måledata er motiveret af resultater fra CS, men der er endnu ingen teoretiske garantier for at opnå nøjagtig CT-rekonstruktion med sparse rekonstruktionsmetoder. Gennem simuleringsstudier påvises eksistensen af en relation mellem billedsparsitet og den krævede mængde måledata i CT. Dette muliggør kvantificering af den reduktion af måledata, som sparse rekonstruktionsmetoder muliggør.

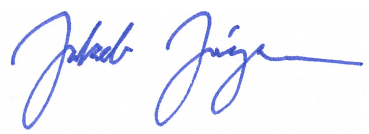
Både det introducerede framework af optimeringsalgoritmer og karakteriseringsmetoderne bidrager til forståelsen af anvendelsesmulighederne for sparse rekonstruktionsmetoder i CT og fungerer som indledende bidrag til et generelt arsenal af beregningsmæssige analyseværktøjer. Således medvirker denne afhandling til at fremme brugen af sparse rekonstruktionsmetoder i praktiske anvendelser af tomografisk rekonstruktion, eksempelvis til lav-dosis CT-scanning.

Preface

This thesis was prepared in partial fulfillment of the requirements for acquiring the PhD degree at the Technical University of Denmark (DTU). The work was carried out between October 2009 and April 2013 in the Section for Scientific Computing, Department of Applied Mathematics and Computer Science (formerly Department of Informatics and Mathematical Modeling), DTU, under supervision of Professor Per Christian Hansen. A significant part of the work was done during two research stays in 2011 and 2012 at the Department of Radiology, University of Chicago, with co-supervisor Associate Professor Emil Y. Sidky. Furthermore, Senior Scientist Søren Schmidt, Department of Physics, DTU, was co-supervisor on the project.

This work was funded in part by the Danish Research Council for Technology and Production Sciences through the project CSI: Computational Science in Imaging (Grant 274-07-0065), and in part by Department of Applied Mathematics and Computer Science (formerly Department of Informatics and Mathematical Modeling). Additional support by the Danish Ministry of Science, Innovation and Higher Education's Elite Research Scholarship is gratefully acknowledged.

Kongens Lyngby, April 11, 2013



Jakob Sauer Jørgensen

Papers included in the thesis

Before 2013, papers were published using birth name, Jakob Heide Jørgensen; from 2013, using married name, Jakob Sauer Jørgensen.

Journal papers

[A] **J. S. Jørgensen**, E. Y. Sidky and X. Pan. *Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in x-ray CT.* IEEE Trans. Med. Imaging, vol. 32, issue 2, pp. 460–473, 2013.

[B] E. Y. Sidky, **J. S. Jørgensen** and X. Pan. *First-order convex feasibility algorithms for x-ray CT.* Med. Phys., vol. 40, issue 3, p. 031115, 2013.

[C] **J. S. Jørgensen**, E. Y. Sidky, P. C. Hansen and X. Pan. *Quantitative study of undersampled recoverability for sparse images in computed tomography.* Submitted to SIAM J. Sci. Comput., 2013.

[D] P. A. Wolf, **J. H. Jørgensen**, T. G. Schmidt and E. Y. Sidky. *Few-view single photon emission computed tomography (SPECT) reconstruction based on a blurred piecewise constant object model.* Submitted to Phys. Med. Biol., 2012.

[E] E. Y. Sidky, **J. H. Jørgensen** and X. Pan. *Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm.* Phys. Med. Biol., vol. 57, issue 10, pp. 3065–3091, 2012.

[F] T. L. Jensen, **J. H. Jørgensen**, P. C. Hansen and S. H. Jensen. *Implementation of an optimal first-order method for strongly convex total variation regularization.* BIT Numer. Math., vol. 52, issue 2, pp. 329–356, 2012.

Conference papers, peer-reviewed

[G] **J. S. Jørgensen**, E. Y. Sidky and X. Pan. *Connecting image sparsity and sampling in iterative reconstruction for limited angle X-ray CT.* Accepted for the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Lake Tahoe, CA, United States, 2013.

[H] E. Y. Sidky, R. Chartrand, **J. S. Jørgensen** and X. Pan. *Nonconvex optimization for improved exploitation of gradient sparsity in CT image reconstruction.* Accepted for the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Lake Tahoe, CA, United States, 2013.

[I] E. Y. Sidky, **J. H. Jørgensen** and X. Pan. *Sampling conditions for gradient-magnitude sparsity based image reconstruction algorithms.* In Medical Imaging 2012: Physics of Medical Imaging, editors N. J. Pelc, R. M. Nishikawa and B. R. Whiting, Proc. of SPIE, vol. 8313, p. 831337, San Diego, CA, United States, 2012.

[J] **J. H. Jørgensen**, E. Y. Sidky and X. Pan. *Ensuring convergence in total-variation-based reconstruction for accurate microcalcification imaging in breast X-ray CT.* In Proceedings of the 2011 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), pp. 2640–2643, Valencia, Spain, 2011.

[K] **J. H. Jørgensen**, T. L. Jensen, P. C. Hansen, S. H. Jensen, E. Y. Sidky and X. Pan. *Accelerated gradient methods for total-variation-based CT image reconstruction.* In Proceedings of the 11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 435–438, Potsdam, Germany, 2011.

[L] **J. H. Jørgensen**, P. C. Hansen, E. Y. Sidky, I. S. Reiser and X. Pan. *Toward optimal X-ray flux utilization in breast CT.* In Proceedings of the 11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 359–362, Potsdam, Germany, 2011.

Software

[S1] T. L. Jensen, **J. H. Jørgensen**, P. C. Hansen and S. H. Jensen. `TVReg`: MATLAB® software package with accelerated first-order optimization methods for total variation regularization, designed for 3D tomographic reconstruction. Accompanying software to paper F. Available from: `www.imm.dtu.dk/~pcha/TVReg`, 2011.

[S2] P. C. Hansen and M. Saxild-Hansen (and **J. H. Jørgensen**). `AIRtools`: MATLAB® software package with implementations of algebraic reconstruction methods and tomographic reconstruction test problems. Accompanying software to [71]. Contributed parallel-beam and fan-beam CT test problem implementations. Available from: `www.imm.dtu.dk/~pcha/AIRtools`, 2011.

[S3] **J. H. Jørgensen**. `tomobox`: MATLAB® software for numerical simulation of 3D computed tomography. Available from: `www.mathworks.com/matlabcentral/fileexchange/28496-tomobox`, 2010.

Other papers not included in thesis

[P1] P. A. Wolf, **J. H. Jørgensen**, T. G. Schmidt and E. Y. Sidky. *A first-order primal-dual reconstruction algorithm for few-view SPECT*. To appear in Proceedings of the 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Anaheim, CA, United States, 2012.

[P2] E. Y. Sidky, **J. H. Jørgensen** and X. Pan. *Characterizing a discrete-to-discrete X-ray transform for iterative image reconstruction with limited angular-range scanning in CT*. To appear in Proceedings of the 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Anaheim, CA, United States, 2012.

[P3] E. Y. Sidky, **J. H. Jørgensen** and X. Pan. *Convergence of iterative image reconstruction algorithms for digital breast tomosynthesis*. To appear in Proceedings of the 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Anaheim, CA, United States, 2012.

[P4] **J. H. Jørgensen**, E. Y. Sidky and X. Pan. *Toward quantifying admissible undersampling of sparsity-exploiting iterative image reconstruction for X-ray CT*. In Proceedings of the Second International Conference on Image Formation in X-Ray Computed Tomography, pp. 161–164, Salt Lake City, UT, United States, 2012.

[P5] E. Y. Sidky, **J. H. Jørgensen** and X. Pan. *Convex optimization prototyping for iterative image reconstruction in X-ray CT*. In Proceedings of the Second International Conference on Image Formation in X-Ray Computed Tomography, pp. 343–347, Salt Lake City, UT, United States, 2012.

[P6] P. C. Hansen and **J. H. Jørgensen**. *Total variation and tomographic imaging from projections*. In Proceedings of the 36th Conference of the Dutch-Flemish Numerical Analysis Communities (WSC 2011), pp. 44–51, Zeist, The Netherlands, 2011, invited contribution.

Acknowledgments

I wish to express my gratitude to a number of people, without whom this thesis would not have been possible. First, I would like to thank my supervisor Per Christian Hansen for introducing me to the world of inverse problems and tomography and for uncountable inspiring discussions. I want to thank my co-supervisor Søren Schmidt for sharing his knowledge on practical aspects of tomographic reconstruction. I am grateful to Anders Skajaa and Martin Skovgaard Andersen for all their insights on optimization and algorithms. Also, I would like to thank all my fellow graduate students for some very enjoyable years in the Section for Scientific Computing. Big thanks to Tobias Lindstrøm Jensen for a wonderful collaboration in the CSI project, and to Søren Holdt Jensen, for hosting me during visits to Aalborg University.

I am extremely thankful to my co-supervisor Emil Sidky and Prof. Xiaochuan Pan for welcoming me into their University of Chicago lab and into their lives. My two Chicago stays have been truly rewarding both personally and on theoretical, computational and practical aspects of CT image reconstruction. The graduate students at the Department of Radiology also deserve to be thanked for making my time in Chicago absolutely amazing.

I also wish to thank a number of collaborators and colleagues, including Christian Kruschel, Dirk Lorenz, Judit Chamorro, Rick Chartrand, Paul Wolf, Taly Gilat Schmidt, Ingrid Reiser and Klaus Mosegaard for many inspiring discussions about tomographic reconstruction, algorithms and theoretical aspects.

Finally, I would like to thank my wife Cathrine and son Viktor as well as the rest of my family and my friends for all their love, support and understanding.

List of Symbols

A	System matrix: discrete-to-discrete CT forward operator.
a_i	Row or column of system matrix A, depending on context.
$a_{i,j}$	Entry (i, j) in system matrix: path length of x-ray i through pixel j.
b	Discrete-domain sinogram.
b_i	Discrete-domain sinogram value, indexed by i.
CC	Continuous-to-continuous imaging model.
CD	Continuous-to-discrete imaging model.
DC	Discrete-to-continuous imaging model.
DD	Discrete-to-discrete imaging model.
D	Finite-difference approximation of derivative operator.
D_j	Finite-difference approximation of derivative operator at pixel j.
$\delta(\cdot)$	The Dirac delta function.
δ_s	Restricted isometry property constant of order s.
Δx	Side length of a pixel.
e	Discrete noise signal.
ϵ	Regularization parameter in data-constrained formulation.
$\mathcal{F}_1$	Fourier transform in 1D.
$F^\star$	Optimal value of objective function $F(u)$ at optimal solution $u^\star$.
$F(u)$	Objective function of optimization problem in u.
$f(\mathbf{x})$	Continuous-domain image.
$g(\mathbf{y})$	Continuous-domain sinogram.
$g(\mathbf{y}_L)$	Continuous-domain sinogram value corresponding to line L.
$\mathcal{H}_{\mathbb{L}}$	Continuous-to-continuous (CC) transform for line set $\mathbb{L}$.
$\mathcal{H}_{\mathbb{L}_{\text{fin}}}$	Continuous-to-discrete (CD) transform for finite line set $\mathbb{L}_{\text{fin}}$.
(h, ℓ)	Double pixel index.
I_0	X-ray intensity before passing through image domain.
I_L	X-ray intensity after passing through image domain.

i	Discrete-domain sinogram index.
$\hat{\imath}$	The imaginary unit $\sqrt{-1}$.
j	Single pixel index.
$J_{\mathrm{TV}} f$	Continuous-domain total variation functional on image $f(\mathbf{x})$.
k	Iteration counter.
$\kappa(A)$	Condition number of matrix A.
L	Line followed by an x-ray through the image domain.
λ	Regularization parameter in regularized formulation.
L_i	Line followed by an x-ray through the image domain, indexed by i.
$\mathbb{L}$	Set of lines passing through the image domain.
$\mathbb{L}_{\mathrm{fin}}$	A finite set of lines passing through the image domain.
M	Discrete-domain number of measurements.
$\mu(A)$	Coherence of matrix A.
$\mu(\Phi, \Psi)$	Mutual coherence of pair of matrices Φ, Ψ.
n	Dimension of continuous-domain image, e.g., $n = 2$ for a 2D image.
N	Discrete-domain number of pixels/expansion functions.
N_b	Number of detector elements in a discrete projection.
N_{s}	Number of pixels in each dimension of discrete image.
N_v	Number of projections.
Ω	Continuous-domain image domain.
Φ	Compressed sensing orthonormal sensing matrix.
φ	Angular parameter of Radon transform.
φ_v	Projection angle indexed by v.
$P_{h,\ell}(\mathbf{x})$	Pixel expansion function using double pixel index (h, ℓ).
(P_p)	Optimization problem: $\min_u \|u\|_p$ s.t. $Au = b$ for $p = 0, 1, 2$.
$p_j(\mathbf{x})$	Pixel expansion function using single pixel index j.
$p_\varphi(\rho)$	Continuous-domain projection or view (for fixed value of φ).
$\pi_{\mathrm{data}}(b\|u)$	Likelihood function for observing b given u.
$\pi_{\mathrm{post}}(u\|b)$	Posterior distribution for u given observed data b.
$\pi_{\mathrm{prior}}(u)$	Prior distribution on u.
Ψ	Compressed sensing orthonormal representation basis matrix.
$\mathcal{R}$	The Radon transform.
ρ	Line parameter of Radon transform.
ρ_w	Detector element indexed by w.
$R_{\mathrm{TV}}(u)$	Discrete-domain total variation regularizer of image u.
$R(u)$	General regularizer.
S	Matrix in regularizer.
σ_e^2	Variance of distribution for e.
τ	Regularization parameter in regularizer-constrained formulation.
$T(u)$	General data fidelity for image u.
U	Discrete-domain image coefficient array, double-indexed by (h, ℓ).
u	Discrete-domain image coefficient vector, single-indexed by j.
$u^\star$	Optimal solution to optimization problem in u.
$\mathbf{x}$	Continuous-domain image coordinate vector.

x_1, x_2	Continuous-domain image coordinates.
$\mathbf{y}$	Continuous-domain sinogram coordinate vector.

Contents

Chapter 1

Introduction

Computed tomography (CT) is the mathematical technique of reconstructing an image of an object from measurements of its projections. Many applications rely on CT; an important and well-known example is the medical CT-scanner, which is also the focus of the present thesis, but many others exist in areas such as biomedical imaging, materials science and geophysics. Even though the CT-scanner has been established as an indispensable medical imaging tool for decades, it is still subject to active research. The CT-scanners of today are designed for classical analytical reconstruction methods such as filtered back-projection (FBP). The achievable image quality is directly related to the x-ray dose given to the patient, and in order to obtain an image of sufficiently high quality, a relatively high x-ray dose must be used.

1.1 Low-dose CT by sparse reconstruction

Motivated by an increasing focus on the potentially harmful effects of CT-scans, a recent trend in CT research has been to develop low-dose imaging techniques. Low-dose imaging is relevant in diagnostic CT scanners to reduce the accumulated exposure that a patient is subjected to through a series of scans associated with a treatment. Low-dose CT imaging can potentially also enable new applications that are currently prevented by the high dose levels needed. For instance,

a dedicated breast CT scanner is being developed with the intention to supplement mammography in periodic screening for breast cancer. Operating in an application, in which a large population fraction will routinely be exposed to x-ray radiation, puts strict limitations on the allowable dose. Low-dose imaging is also relevant in other applications of CT, for example, in biomedical imaging and materials science to prevent causing damage to the subject under study.

A main driving factor for the potential for low-dose imaging has been the emergence of sparse reconstruction methods, proposed for example in compressed sensing (CS). Sparse reconstruction enables accurate reconstruction from a reduced number of measurements under the assumption of a sparse image and certain restrictions on the measurement process. The field of sparse reconstruction has seen tremendous development over the past decade or so. Theoretical results show great promise for achieving accurate reconstruction from heavily undersampled data. At the same time, a lot of effort has been devoted into developing fast algorithms for variational image reconstruction in general, and sparse image reconstruction in particular. Similarly, a multitude of reconstruction formulations exploiting sparsity in many different ways have been proposed. The potential for successful application to CT has been demonstrated empirically in a number of studies both in simulation and applied to real data.

1.2 Aims of the thesis

This thesis is motivated by a desire to *understand* and *quantify* the factors that determine the attainable reconstruction quality by sparse reconstruction methods in CT. In other words, the aim is to *characterize* the use of sparse reconstruction methods in CT. Sparse reconstruction is approaching a stage where a plethora of formulations and algorithms are available and the initial proof-of-concept has been established. As in [93], one can ask: *What major factors prevent sparse reconstruction from transitioning into routine use in applications such as CT?* We argue that, for this purpose, the most critical aspects are not to construct additional algorithms competing to be slightly faster than existing ones, and similarly, not to come up with new variants of sparse reconstruction formulations. These are indeed important research directions to pursue, but in our opinion they are unlikely to change the state-of-the-art in CT, before more fundamental questions have been answered. Rather, we find that an improved understanding of the practical potential is needed; in particular a quantitative understanding of the factors that determine the reconstruction quality.

Imaging through CT scanning is a complex subject that can be considered a chain with many different interacting phases. From design and manufacturing

of the scanner, over the data acquisition protocol, data preprocessing, and the image reconstruction procedure, to postprocessing, image analysis, evaluation and decision-making based on the reconstructed image. Thus, CT scanning is a very applied and practical subject, and as such, it is natural that a large body of CT imaging research considers this entire imaging chain. It is attempted to develop practical, procedural steps that lead to an empirically observable improvement on solving the given imaging task, for example detection of malignant tumors. With the enormous number of design choices on everything including acquisition, reconstruction algorithms and quality assessment methods it can be difficult to acquire a complete understanding of all individual steps in the chain, and hence determine in which way the final output depends on each step.

A very different research approach consists of focusing on a single step or even sub-step of the imaging chain and study that problem in depth. The applied research goal of demonstrating better utility of a full procedure is then replaced by more fundamental research goals of providing better understanding of each individual step and establishing relations between parameters of interest. Hopefully, the gathered insights will then help applied CT researchers devise useful novel imaging modalities.

The present thesis takes the latter approach and focuses on the image reconstruction block of the imaging chain. The aim is not to propose and claim superiority of new reconstruction methods. In contrast, a number of sub-problems of image reconstruction are studied more in detail and several tools are developed for improving the understanding of which factors affect the reconstruction quality that can be obtained by sparse reconstruction methods.

We envision a "computational toolbox" of methods for characterizing sparse reconstruction methods: Systematic ways of obtaining fundamental, quantitative insight into capabilities of sparse reconstruction methods. In the thesis we take the initial steps toward such characterization methods by proposing several tools and outline a number of future paths to pursue.

The thesis contributions fall in two major categories:

1. Development of "prototyping" optimization algorithms and software to enable seamless experimentation with sparse reconstruction methods based on different optimization problems.
2. Development of characterization tools for establishing quantitative understanding of reconstruction quality attainable by sparse reconstruction methods in CT.

Applying sparse reconstruction methods, as proposed in the mathematical imaging community, to the practical application of CT is very much a translational

research effort. It requires thorough understanding of theoretical, computational and application-oriented matters. It is our impression that only few researchers are truly interested in pursuing this translational challenge. With this thesis, we are doing precisely that, and we note that the thesis contributions have been published in mathematical/numerical as well as application-oriented venues.

1.3 Structure of the thesis

This thesis is structured in two parts. The first part sets the stage and provides the reader with background knowledge of inverse problems and CT imaging; this part can be skimmed on a first reading by readers primarily interested in the applications and the results, which are described in the second part.

In chapter 2 we cover the fundamentals of CT, including the physical set-up, standard configurations, different imaging models used and an overview of classical analytical and algebraic reconstruction methods. In chapter 3 we describe how CT fits into the general framework of inverse problems. We give a general presentation of regularization methods to overcome the challenge of solving inverse problems. Particular emphasis is put on total variation regularization and we discuss aspects of optimization and numerical algorithms relevant for regularization. We present the application to CT motivated by low-dose imaging. In chapter 4, we give a general presentation of sparse image reconstruction methods. We cover several reconstruction formulations and algorithms, theoretical guarantees of accurate reconstruction and address the application to CT.

In chapter 5 we outline some remaining challenges for sparse reconstruction in CT. In particular, we identify a need for improving the fundamental quantitative understanding of factors that affect the achievable image quality. We suggest to develop a set of characterization tools for systematically establishing such quantitative understanding. We also develop a general framework for analyzing sparse reconstruction in CT. The framework, which is explained in section 5.2, puts most of the background material into context and it can therefore be useful to keep an eye on while reading the background chapters. In chapter 6 we describe the thesis contributions in the setting of developing characterization tools for sparse image reconstruction methods in CT, while referring to the relevant papers in the appendix. As mentioned in the previous section, the contributions can be broadly split into development of prototyping optimization algorithms and software and development of computational characterization tools, for example to quantitatively establish the effect of image sparsity on reconstruction quality. In chapter 7 we discuss the obtained results and outline future directions, before concluding the thesis.

Chapter 2

Computed tomography

In this chapter we give an introduction to computed tomography. We cover the most important historical developments, some standard configurations and the underlying physical model. We describe the relevant imaging models and reconstruction methods used in the two classical regimes of reconstruction: analytical and algebraic reconstruction. The focus is on medical CT but most of the material, for example the imaging models and reconstruction methods, is relevant for general tomographic imaging.

2.1 Tomographic imaging

Computed tomography (CT), or tomographic imaging, is to determine an image of an object from measurements of its projections. CT is used in numerous applications, where we are interested in looking at something that we do not have direct access to, for example the interior of the human body. Instead of having to physically "open up" the object, we can acquire projections of the object from the outside, and then through CT obtain an image of the inside. One of the most well-known examples, and the focus of the present thesis, is the medical CT-scanner, which acquires projection images of a patient using x-rays. Other examples abound, in the medical setting we have positron emission

tomography (PET), single photon emission computed tomography (SPECT), magnetic resonance imaging (MRI); and other examples include use in materials science (for example to monitor the microstructure of metals), geoscience (to find oil or groundwater) and astronomy (to study properties of distant stars or planets).

The mathematics underlying tomographic imaging is described in numerous works; we mention a few of the "standard references", [4, 20, 56, 72, 78, 88, 89], all of which have a medical imaging focus. The focus in the work done for this thesis has been on the tomographic imaging in the setting of medical CT. Therefore the specific results obtained are valid for this particular application, however, the general ideas and proposed methodologies are not limited to medical CT. We foresee that similar insights can be obtained for other applications of tomographic imaging.

2.1.1 Design and history of the CT scanner

The history of the medical CT scanner is described, e.g., in the review articles [10, 44, 79, 94]. We give a brief recap. The physical foundation of the CT scanner was provided by Röntgen, who discovered x-rays in 1895, for which he received the first Nobel prize in physics in 1901 [1]. X-ray imaging was quickly developed and provided a unique noninvasive way to look at the interior of a patient for diagnostic purposes. An inherent problem was the lack of depth information. To address this problem, Hounsfield developed the first CT scanner in the early 1970'es [74]. With help from Cormack, the CT scanner was successfully put into clinical use, and Hounsfield and Cormack shared the Nobel prize in Medicine in 1979 [1]. While the first CT scanner was designed for head scanning and took hours to acquire sufficient data and compute a reconstructed image, development quickly took place and resulted in full-body and other dedicated scanners as well as much faster data acquisition and reconstruction.

The first generation of scanners used a parallel-beam geometry, see Figure 2.1, in which a single x-ray source and detector element were used to record a single data point at a time. This type of scanner was quickly replaced by a divergent beam, also called a fan-beam geometry, see Figure 2.1, using a detector with a curved array of detector elements, which could acquire all the data points in a projection view at a time, thus considerably reducing the data acquisition time. With the parallel-beam and fan-beam configurations, 3D imaging was possible only through multiplanar 2D, i.e., piecing together individual 2D reconstructions. Fully 3D image reconstruction became possible with the development of scanners based on a cone-beam geometry, i.e., a 3D divergent-beam equivalent of the 2D fan-beam geometry.

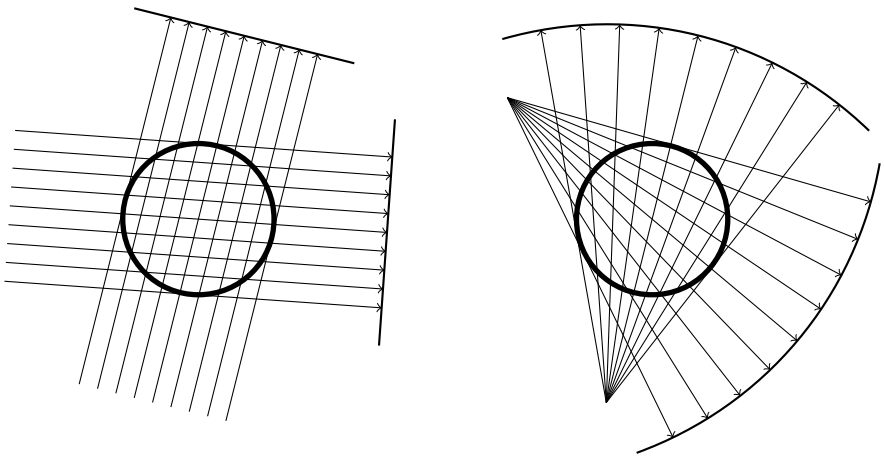


Figure 2.1: Left: the parallel-beam CT geometry. Right: the fan-beam geometry.

Today, a variety of different medical scanning modalities based on the CT imaging principles exist, including diagnostic CT, tomosynthesis, dedicated breast CT, micro-CT, dental CT, dynamic CT and phase-contrast CT. The reconstruction principle is shared but there are large differences in physical design. While in a conventional CT scanner, projections are acquired from 360° around the patient, in tomosynthesis, for example, projections can only be acquired from a restricted set of angles, leading to the so-called limited angle tomography problem. As one might expect, this makes the reconstruction problem more difficult.

2.1.2 The underlying physics

As mentioned, tomographic imaging amounts to reconstruction from projections. In the setting of medical CT, projections are obtained by passing x-rays through the patient and measuring how much the x-rays are attenuated. The x-ray attenuation in tissue primarily depends on the tissue density. Each type of tissue has an associated attenuation coefficient and the goal of CT is to determine the attenuation coefficient across the object. X-ray attenuation in tissue can be described by Lambert-Beer's law, see e.g. [20]. If $f(\mathbf{x})$ is the attenuation coefficient at the physical position $\mathbf{x}$ in the object, L is a line through the object, and I_0 and I_L are x-ray intensities before and after the object, then we have

$$I_L = I_0 \exp\left(-\int_L f(\mathbf{x}) d\mathbf{x}\right). \tag{2.1}$$

That is, the x-ray intensity after passing through the domain has been reduced by a factor given by the line integral of the attenuation coefficient along the line L. This model assumes that the x-rays have a single wavelength and the attenuation coefficient is specific for this wavelength. This assumption is appropriate for synchrotron imaging where x-rays can be made to have a single specific wavelength. A conventional medical CT scanner uses an x-ray tube and generates a broader spectrum. A more accurate model can be set up to account for this, but in practice Lambert-Beer's law is often used despite the fact that more wavelengths are present in the x-ray beam.

Lambert-Beer's law can be considered an average-case model of the behavior of x-rays passing through the object [20, 58]. In practice, some statistical fluctuations will be present in the number of photons emitted from the x-ray source and in the number recorded by the detector. It can be shown that ideally the detector counts follow a Poisson distribution. However, other types of noise and inconsistencies such as scatter, beam-hardening and electronic noise contribute to making the observed inconsistencies with respect to Lambert-Beer's model non-Poissonian. Further, the data is log-transformed to obtain projection data in (2.2), which further complicates the noise distribution. In practice, due to the many unknown factors, it is common to assume a Gaussian noise-model, which we discuss in subsection 3.1.4.

The quality of the acquired data is closely connected to the x-ray dose the patient is given. Dose is a complex subject and we will only touch it briefly. The variance of the data is inversely proportional to the x-ray dose given to the patient, that is, if the dose is reduced, then the data variance increases, corresponding to a higher noise level [17, 20]. It is clear from this, that if the x-ray intensity is reduced, then the data quality becomes poorer.

2.2 Imaging models

In this section we describe the two classic regimes of CT image reconstruction: the analytical methods, which are based on a continuous formulation, and the algebraic methods, which are based on a discretized formulation. While the imaging models have similarities, there are also important differences. In the literature, it is often not stated which imaging model is used in a particular work, which can lead to confusion.

The continuous-to-continuous (CC) imaging model is the fundamental model with which it is possible to study many important questions such as existence, uniqueness and stability of a solution. When it comes to applying the model

to real data, we need to account for the finite set of line integral measurements that are acquired by a scanner. By doing so we obtain a continuous-to-discrete (CD) imaging model. Analytical reconstruction methods, see subsection 2.3.1, are based on (possibly approximately) inverting the CD imaging model, which means that they produce a continuous image. This is in contrast to algebraic methods, see subsection 2.3.2, which are based on inverting the discrete-to-discrete (DD) imaging model, and therefore produce a discrete image.

2.2.1 The continuous-to-continuous imaging model

Following [4], we are interested in obtaining an image of a physical object $f(\mathbf{x})$ that lives in the continuous domain. For simplicity of the presentation, we assume here a two-dimensional object, i.e., $\mathbf{x} \in \mathbb{R}^2$. We assume the support of $f(\mathbf{x})$ is contained within a disk.

We begin from Lambert-Beer's law (2.1). By taking the logarithm we obtain

$$\int_L f(\mathbf{x}) d\mathbf{x} = \log \frac{I_0}{I_L} = g(\mathbf{y}_L). \tag{2.2}$$

A measurement I_L, before taking the logarithm, is called a transmission measurement, while a measurement $g(\mathbf{y}_L)$ is called a projection measurement. Let $\mathbb{L}$ denote a set of lines L. Given an object $f(\mathbf{x})$ and a set of lines $\mathbb{L}$, the *forward problem* consists of computing all the right-hand sides $g(\mathbf{y})$, where $\mathbf{y} \in \mathbb{R}^2$ parametrizes the lines in $\mathbb{L}$. We can write this CC imaging model as

$$\mathcal{H}_{\mathbb{L}} f(\mathbf{x}) = g(\mathbf{y}), \tag{2.3}$$

where $\mathcal{H}_{\mathbb{L}}$ denotes the CC transform operator. The *inverse problem* arises when instead of $f(\mathbf{x})$ we are given $g(\mathbf{y})$ and we want to reconstruct $f(\mathbf{x})$.

The central example of a CC transform is the Radon transform $\mathcal{R}$, which arises when taking the complete set of all possible lines passing through the object. The transform is named after Johann Radon, who laid the foundation of CT imaging in his seminal 1917-paper, [99] in German, see also an English translation [100], by proving that an object is uniquely determined by its Radon transform. A set of Radon transform data is called a sinogram.

For writing the Radon transform explicitly, the lines are commonly parametrized using an angular parameter $\varphi \in [0, \pi[$ and a line parameter $\rho \in \mathbb{R}$, see Figure 2.2,

$$\rho = x_1 \cos\varphi + x_2 \sin\varphi. \tag{2.4}$$

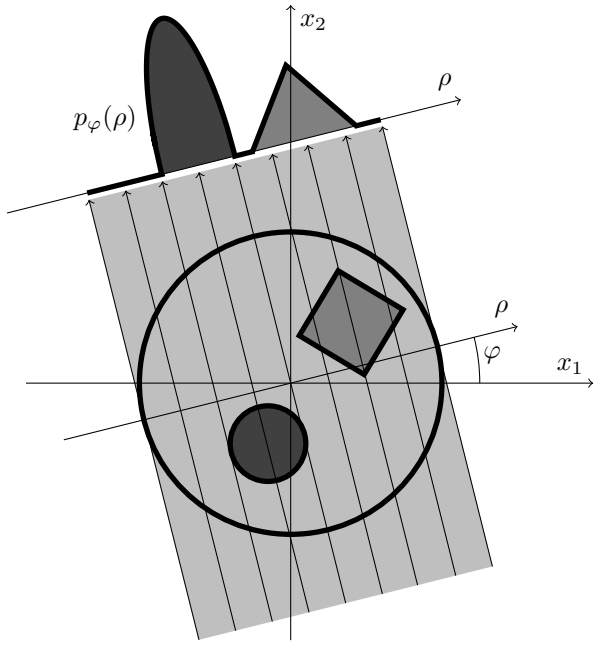


Figure 2.2: The Radon transform in 2D. A projection at a single angle φ is shown. The object's attenuation coefficient in the dark gray disk is twice of that in the light gray square, causing the projection of the disk to have a twice as large maximal value as that of the square.

The Radon transform can then be written as

$$[\mathcal{R}f](\rho,\varphi) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1,x_2)\delta(\rho - x_1\cos\varphi - x_2\sin\varphi)dx_1dx_2, \tag{2.5}$$

where $\delta(\cdot)$ is the Dirac delta function. The values of the Radon transform for a constant angular parameter φ_0 is called a projection or a view, written

$$p_{\varphi_0}(\rho) = [\mathcal{R}f](\rho,\varphi_0). \tag{2.6}$$

The Radon transform corresponds to the parallel-beam geometry and is illustrated in Figure 2.2. A single projection of the object with a square and a disk-shaped feature is shown.

More generally, in $\mathbb{R}^n$ the Radon transform consists of integrating the object over all $(n-1)$-dimensional hyperplanes; in 3D, for example, over all planes. Many other integral transforms are important to CT and medical imaging, for example the related *x-ray transform* that integrates along lines instead of hyperplanes.

2.2.2 The continuous-to-discrete imaging model

We will mainly consider a simple method for data-space discretization, namely to select a finite number M of lines $\{L_i\}_{i=1,\dots,M}$ from the full set of lines $\mathbb{L}$. This corresponds to an assumption of the source as well as the detector elements having zero width. A more accurate discretization method would take into account the detector-element width, but it is common to use the simpler model and we stick with this choice. We denote the finite set of lines $\mathbb{L}_{\text{fin}}$. The finite set of data samples is collected in a vector b, where

$$b_i = g(\mathbf{y}_{L_i}) = \log \frac{I_0}{I_{L_i}}, \quad i = 1, \dots, M. \tag{2.7}$$

We can then write the CD imaging model as

$$\mathcal{H}_{\mathbb{L}_{\text{fin}}} f(\mathbf{x}) = b, \tag{2.8}$$

where $\mathcal{H}_{\mathbb{L}_{\text{fin}}}$ is a continuous-to-discrete operator. Which lines that are selected rely on the scanner geometry. For 2D fan-beam and parallel-beam configurations, we could for example select a discrete set of projection angles $\{\varphi_v\}_{v=1,\dots,N_v}$, as well as a discrete set of detector positions $\{\rho_w\}_{w=1,\dots,N_b}$ with $M = N_v N_b$ and use the corresponding set of lines.

2.2.3 The discrete-to-discrete imaging model

To derive the discrete-to-discrete (DD) imaging model, we proceed from the CD model by discretizing also the object space. To simplify the presentation we assume a 2D object $f(\mathbf{x})$ over a square Ω of side length 1, $\mathbf{x} \in \Omega = [-1/2, 1/2]^2$. We use in the continuous object space the inner product

$$\langle f_1, f_2 \rangle = \int f_1(\mathbf{x}) f_2(\mathbf{x}) d\mathbf{x}. \tag{2.9}$$

We obtain a discrete representation of the object by expansion in terms of a certain set of expansion functions. In CT, many different expansion functions have been considered, including pixels and their 3D counterpart voxels, as well as so-called blobs [80] and natural pixels [19]. Here we assume pixel expansion functions, obtained by dividing the object space into $N_{\text{s}} \times N_{\text{s}}$ pixels, each of side length $\Delta x = 1/N_{\text{s}}$. The (h, ℓ)th pixel expansion function is supported precisely within pixel (h, ℓ) and is defined by

$$P_{h,\ell}(\mathbf{x}) = \begin{cases} N_{\text{s}}^2 & \text{if } \mathbf{x} \in [x_1^{(h-1)}, x_1^{(h)}] \times [x_2^{(\ell-1)}, x_2^{(\ell)}], \\ 0 & \text{else,} \end{cases} \quad h, \ell = 1, \dots, N_{\text{s}}, \tag{2.10}$$

where the pixel boundaries are given by

$$x_1^{(h)} = -\frac{1}{2} + h\Delta x, \quad h = 0, 1, \ldots, N_\mathrm{s}, \tag{2.11a}$$

$$x_2^{(\ell)} = -\frac{1}{2} + \ell\Delta x, \quad \ell = 0, 1, \ldots, N_\mathrm{s}. \tag{2.11b}$$

The pixel expansion functions are orthogonal due to non-overlapping support, but not scaled to unit norm, since

$$\|P_{h,\ell}\|_2^2 = \langle P_{h,\ell}, P_{h,\ell}\rangle = \int_{x_2^{(\ell-1)}}^{x_2^{(\ell)}} \int_{x_1^{(h-1)}}^{x_1^{(h)}} N_\mathrm{s}^4 dx_1 dx_2 = N_\mathrm{s}^2, \quad h, \ell = 1, \ldots, N_\mathrm{s}. \tag{2.12}$$

We will return to explaining this choice. We want an expansion of $f(\mathbf{x})$ in the pixel expansion functions and due to the non-unit norm of the basis function we include a normalization factor in the sum:

$$f(\mathbf{x}) = \frac{1}{N_\mathrm{s}^2} \sum_{h,\ell=1}^{N_\mathrm{s}} U_{h,\ell} P_{h,\ell}(\mathbf{x}), \tag{2.13}$$

where the coefficients $\{U_{h,\ell}\}_{h,\ell=1,2,\ldots,N_\mathrm{s}}$ specify the discrete representation U, i.e., an array of pixel values. Of course, this expansion is only valid with equality for objects $f(\mathbf{x})$ that are already constant within each pixel; for other objects, the expansion provides an approximation. We can derive a computational expression for $U_{h,\ell}$ by taking inner products with $P_{h,\ell}(\mathbf{x})$ and use orthogonality,

$$\langle f, P_{h,\ell}\rangle = \left\langle \frac{1}{N_\mathrm{s}^2} \sum_{\tilde{h},\tilde{\ell}=1}^{N_\mathrm{s}} U_{\tilde{h},\tilde{\ell}} P_{\tilde{h},\tilde{\ell}}, P_{h,\ell} \right\rangle = \frac{1}{N_\mathrm{s}^2} U_{h,\ell} \langle P_{h,\ell}, P_{h,\ell}\rangle = U_{h,\ell}, \tag{2.14}$$

which means that the (h,ℓ)th pixel value is simply the inner product with the (h,ℓ)th expansion function.

We can now explain our choice of the nonzero value of the expansion function, because this choice makes the pixel value $U_{h,\ell}$ equal to the average value of $f(\mathbf{x})$ over pixel (h,ℓ), for example for an object with constant value f_0 in pixel (h,ℓ) we get the same pixel value:

$$U_{h,\ell} = \int_{x_2^{(\ell-1)}}^{x_2^{(\ell)}} \int_{x_1^{(h-1)}}^{x_1^{(h)}} N_\mathrm{s}^2 f_0 dx_1 dx_2 = N_\mathrm{s}^2 f_0 \frac{1}{N_\mathrm{s}^2} = f_0. \tag{2.15}$$

In some places it will be convenient to replace the double-index notation by a single pixel index. We introduce

$$j = h + (N_\mathrm{s} - 1)\ell, \quad h, \ell = 1, \ldots, N_\mathrm{s}, \quad j = 1, \ldots, N_\mathrm{s}^2. \tag{2.16}$$

In the single index we write the pixel values as the vector $u = \{u_j\}_{1,\dots,N_s^2}$ and the pixel expansion functions as $p_j(\mathbf{x})$, $j = 1, \dots, N_s^2$. Letting $N = N_s^2$ denote the total number of pixels we can also write the expansion in (2.13) simply as

$$f(\mathbf{x}) = \frac{1}{N} \sum_{j=1}^{N} u_j p_j(\mathbf{x}). \tag{2.17}$$

It is easy to use only a sub-set of pixel basis functions from the full square grid. We will often use a disk-shaped region in order to match more closely the disk-shaped object in the continuous domain; more precisely, we take the pixel expansion functions with support inside the largest disk inscribed in the square, see Figure 2.3. In this case we keep letting N denote the actual number of pixels, so (2.17) continues to hold, but N is no longer N_s^2. For the disk-shaped region we have $N \approx (\pi/4)N_s^2$, as computed by the ratio of the disk and square areas, and the approximation becomes better with increasing N_s.

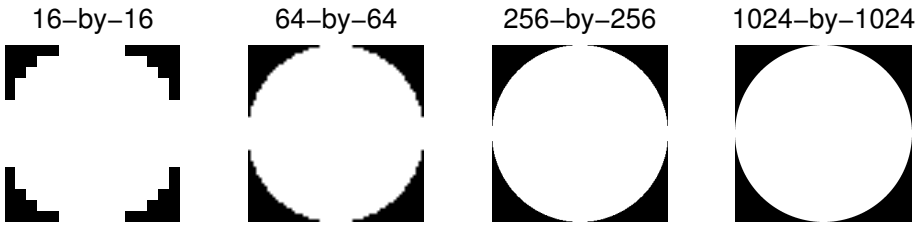


Figure 2.3: Often only pixels within a disk-shaped region in the square domain will be considered. The disk-shaped region is shown as white pixels and the outside as black for four discrete images with different numbers of pixels. When the number of pixels grows, the approximation of a disk becomes better.

With the object expanded in terms of pixel expansion functions we are ready to discretize (2.2) for a given line as indexed by i in (2.7). An example of a discretized object and a given line through it is shown in Figure 2.4. We get

$$\begin{aligned} b_i = g(\mathbf{y}_{L_i}) &= \int_{L_i} f(\mathbf{x}) d\mathbf{x} = \int_{L_i} \frac{1}{N} \sum_{j=1}^{N} u_j p_j(\mathbf{x}) d\mathbf{x} \\ &= \frac{1}{N} \sum_{j=1}^{N} u_j \int_{L_i} p_j(\mathbf{x}) d\mathbf{x}, \quad i = 1, \dots, M. \end{aligned} \tag{2.18}$$

For evaluating the integral, we observe that either the line L_i intersects pixel j or it does not. In the latter case, the integral is zero due to the support of p_j. In the former case, since $p_j(\mathbf{x})$ is constant and equal to N_s^2 inside the pixel, the integral equals $N_s^2 a_{i,j}$, where $a_{i,j}$ is the path length of L_i through pixel j. But

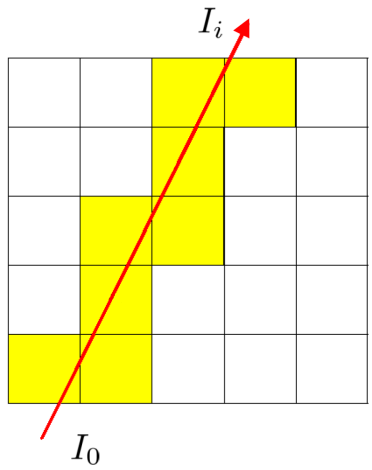


Figure 2.4: A 5×5-pixel example of a discrete image and the path of a single x-ray through it. Each pixel intersected by the ray is yellow and the path lengths inside each of the yellow pixels are the nonzeros of the row of the system matrix corresponding to the shown ray.

if L_i is not intersecting pixel j then the path length is zero, so we can combine the two cases to

$$\int_{L_i} p_j(\mathbf{x}) d\mathbf{x} = N_\mathrm{s}^2 a_{i,j}, \quad j = 1, \ldots, N. \tag{2.19}$$

Using this yields

$$b_i = \frac{1}{N} \sum_{j=1}^{N} u_j N_\mathrm{s}^2 a_{i,j} = \sum_{j=1}^{N} u_j a_{i,j}, \quad i = 1, \ldots, M. \tag{2.20}$$

By setting up the system matrix $A = \{a_{i,j}\}_{i=1,\ldots,M,\ j=1,\ldots,N}$, this system of linear equations can be put in matrix-vector form:

$$Au = b. \tag{2.21}$$

This is the DD imaging model. It is also called the algebraic model. Note that for other expansion functions, each element $a_{i,j}$ of the system matrix can still be computed as the integral along the ith ray with the jth expansion function.

The specific discretization method we used for the data and object spaces is referred to as the line-intersection method, the center line method, and also as Siddon's method, although Siddon did not suggest the method itself but a fast implementation of it [106]. Other methods such as area-weighting in which planar integrals over pixels replace line integrals to account for the nonzero

width of the source and detector bins, ray-tracing with nearest neighbor interpolation and the distance-driven method [43] can be used instead with different advantages and drawbacks.

2.2.4 The discrete-to-continuous imaging model

The fourth and last imaging model is the discrete-to-continuous (DC) imaging model. It is rarely used in practice but in our paper A we use the DC model to study the limiting case of increasing the number of measurements toward infinity while keeping the finite representation fixed. Except for that application of the DC model, we will not use it in the thesis work.

2.3 Reconstruction methods

2.3.1 Analytical reconstruction

Methods for image reconstruction based on analytical transform-inversion are referred to as analytical reconstruction methods or direct reconstruction methods. The idea is to construct an exact, or in some cases approximate, analytical inverse of the CC imaging model of the CT configuration of interest. For reconstruction from actual data, the CC inverse must be discretized in the data domain to obtain an approximate inverse of a CD imaging model. The object space is *not* discretized, which means that an analytical reconstruction is a continuous function, that can in principle be evaluated at all points in the image domain. Typically, only a set of samples is evaluated for displaying the reconstruction, e.g., on a pixel grid, but it is important to keep in mind that the reconstruction is actually in the continuous domain. This is a clear difference from algebraic reconstruction, discussed in subsection 2.3.2, where inherently only an image represented by its finite set of expansion coefficients is reconstructed.

We will only scrape the surface of the massive field of analytical inversion by giving the most common analytical inversion method for the Radon transform, namely the filtered back-projection (FBP) method.

The derivation of the FBP method is straightforward and given in most references on the mathematics of medical imaging, e.g. in [20]; here we give only the resulting analytical inversion formula. We denote the 1D Fourier transform of

a projection $p_\varphi(\rho)$ with respect to ρ as

$$[\mathcal{F}_1 p_\varphi](\omega) = \int_{-\infty}^{\infty} p_\varphi(\rho) e^{-2\pi \hat{\imath} \rho \omega} d\rho, \tag{2.22}$$

where $\hat{\imath}$ denotes the imaginary unit. The image can be reconstructed from parallel-beam projections over an angular range of 180°, or $\varphi \in [0, \pi[$, through the formula

$$f(x_1, x_2) = \int_0^{\pi} \widetilde{p_\varphi}(\rho) d\varphi, \tag{2.23}$$

where $\rho = x_1 \cos\varphi + x_2 \sin\varphi$, and

$$\widetilde{p_\varphi}(\rho) = \int_{-\infty}^{\infty} [\mathcal{F}_1 p_\varphi](\omega) |\omega| e^{2\pi \hat{\imath} \rho \omega} d\omega \tag{2.24}$$

is called a *filtered projection*, since it is obtained by filtering the projection $p_\varphi(\rho)$ with a ramp filter, i.e., with frequency response $|\omega|$. The step in (2.23) is called *back-projection*, as the result for each value of φ to "smear out" its argument over the image domain along lines with normal vector $(\cos\varphi, \sin\varphi)^T$ and through the integration combine the contributions from all $\varphi \in [0, \pi[$.

An alternative analytical inversion formula for the Radon transform is the back-projection filtration (BPF) method [128], in which the order of filtering and back-projection is reversed compared to FBP.

For application to real data, the analytical reconstruction methods must be adapted to the CD imaging model. For example, the CD version of FBP follows by discretization of (2.23) and (2.24). Depending on how the discretization is done, different variants of the CD FBP are obtained. To handle noisy data the algorithm is normally further equipped with an additional filter of the low-pass type, for example a Hamming or Hann filter. This is because (2.24) corresponds to a high-pass filter, which amplifies high-frequency noise.

For analytically inverting the CC Radon transform, all methods (such as FBP and BPF) are equivalent in the sense that they exactly recover the original object $f(\mathbf{x})$ from a complete and noise-free sinogram. However, when adapted to discrete data, differences show up, and especially inconsistent data caused by noise, modeling errors and many other factors can lead to very different reconstructions, thereby raising the question of which method performs better in practice. There is no simple answer, but it is a fact that most commercial CT scanner manufacturers have chosen to employ some form of CD FBP [93]. For 3D circular cone-beam CT, a widely used inversion method was proposed by Feldkamp, Davis and Kress [57], and is now known now as the FDK method. Interestingly, the inverse is only approximate but when applied to clinical data the method is generally considered able to deliver better reconstructions than exact inversion formulas.

2.3.2 Algebraic reconstruction

Methods for inverting the DD imaging model, $Au = b$, are called algebraic reconstruction methods. In the medical CT community they are also referred to as iterative image reconstruction (IIR) methods. Since the DD imaging model amounts to a system of linear equations, we can in principle apply any method for solving linear systems, for example direct methods such as Gaussian elimination. In practice, however, the size of the systems calls for application of an iterative method, which is possibly the cause of the name iterative image reconstruction.

Two classes of iterative methods commonly used in tomographic reconstruction are the algebraic reconstruction techniques (ART) and the simultaneous iterative reconstruction techniques (SIRT). A good overview is given in [71], upon which the following brief presentation is based.

The basic ART method was introduced to the CT community by Gordon, Bender and Herman in 1970 [63] and is equivalent to the Kaczmarz method [76] from 1937. It is a so-called row-action method and consists of the (inner) iteration indexed by the row number i of the system matrix A,

$$u^{(k,i)} = u^{(k-1,i)} + \lambda^{(k)} \frac{b_i - a_i^T u^{(k-1,i)}}{\|a_i\|_2^2} a_i, \ k = 1, 2, \ldots, \ i = 1, 2, \ldots, M, \quad (2.25)$$

where a_i^T denotes the ith row of A. $\lambda^{(k)}$ is called the relaxation parameter. In the Kaczmarz method the rows are traversed in order $i = 1, \ldots, M$ and such a sweep makes up an outer iteration as indexed by k. Other ART variants are obtained by using a different order.

The relaxation parameter must be positive and less than 2 for ART to converge. In case of a consistent linear system, convergence means to a solution of the system. In case of an inconsistent system, convergence means to a "limit-cycle", i.e., the same sequence of points will be cycled. How fast ART converges depends, among other things, on the relaxation parameter. It can be fixed, decay at a predetermined rate or recomputed in various ways in each iteration.

The SIRT methods replace the one-by-one application of rows in ART by a single simultaneous iteration step, in the general notation from [71],

$$u^{(k)} = u^{(k-1)} + \lambda^{(k)} T_1 A^T T_2 (b - A u^{(k)}), \quad (2.26)$$

where T_1 and T_2 are symmetric positive matrices that can be chosen in different ways to obtain different methods. The most basic of them, Landweber's

method, is obtained by setting both to the identity matrix. Other variants include Cimmino's method, component averaging (CAV), diagonally relaxed orthogonal projection (DROP), and the simultaneous algebraic reconstruction technique (SART).

Each iterative method must be equipped with a stopping criterion to determine when a sufficiently accurate solution has been determined. Several stopping criteria are available from the literature including choices that assume knowledge of the noise level in data, such as the discrepancy principle, and others, such as the normalized cumulative periodogram (NCP), see e.g. [69]. It is generally recognized, however, that there is no generally best stopping criterion, so in practice a choice must be made, and the resulting approximate solution will depend on this choice.

2.4 Summary

We have now given an introduction to the field of CT image reconstruction. We presented highlights of the history of the CT scanner, described some standard configurations such as parallel-beam, fan-beam and cone-beam, and we saw that CT imaging is built upon a model of x-ray attenuation in tissue. Further, we described the imaging models used for analytical and algebraic reconstruction, i.e., the continuous-to-discrete (CD) and the discrete-to-discrete (DD) imaging models, respectively, and gave a brief introduction to both types of reconstruction methods.

Chapter 3

Inverse problems and regularization

In this chapter we describe the field of inverse problems, to which CT can be considered to belong. By taking a more general perspective, it is possible to gain significant insights into CT image reconstruction. We give a brief overview of important aspects of inverse problems, including the notions of ill-posed and ill-conditioned problems, the use of regularization to obtain meaningful solutions, and the Bayesian statistical perspective. We focus on total variation (TV) regularization for obtaining reconstructions with sharp edges and describe some advantages and drawbacks of TV-regularization. The motivation for and strategies for application to CT image reconstruction are presented. Finally, we show that introducing regularization amounts to solving optimization problems and we cover some optimization aspects to be aware of, including various properties of optimization problems and algorithmic aspects.

Inverse problems and variational methods can be considered both in the setting of the CC imaging model (2.3) and the DD imaging (or algebraic) model (2.21). For working with and solving inverse problems numerically on a computer, the relevant choice is the DD imaging model, and we will thus restrict our presentation to this model.

3.1 Inverse problems

3.1.1 The forward and inverse problem

Computed tomography is an example of an *inverse problem*. Inverse problems arise whenever we are interested in looking at something that can only be indirectly observed. Assume there is an *object* that cannot be observed directly; in the example of CT, we want to obtain cross-section images or even a full 3D image of the human body. We cannot look inside the body directly. Instead we can record projection images of the body by passing x-rays through it; in general we acquire some *observations* or *data*. We have a physical *model* describing how the object is transformed into the observations; for CT, the model describes the physics of x-ray attenuation in tissue. The task of determining the observations from the object and knowledge of the model is called the *forward problem*. The *inverse problem* is the opposite, namely to reconstruct the object from the observations. In CT, the forward problem corresponds to computing projection images from a known object subject to a given scanning configuration. In practice, the goal in CT is to solve the inverse problem of reconstructing the object from the set of acquired projection images.

The different types of imaging models described in section 2.2 for CT also apply to general inverse problems. The DD imaging model $Au = b$ is also called a linear discrete inverse problem.

3.1.2 Ill-posed and ill-conditioned problems

The linear discrete inverse problem looks innocent as it is simply a system of linear equations but can be very challenging to solve. The problem is that inverse problems are often *ill-posed*. Hadamard [66] gave conditions for a problem to be *well-posed*:

1. **Existence:** The problem must have a solution.
2. **Uniqueness:** The solution must be unique.
3. **Stability:** A small data change must only give a small solution change.

If a problem does not satisfy these conditions, it is called *ill-posed*. Depending on the particular context, the linear discrete inverse problem can fail to satisfy either of the three conditions. For example, assuming that the linear system is consistent, if A has fewer rows than columns, then there is a nontrivial nullspace and hence infinitely many solutions, so the uniqueness condition fails. If instead

the linear system has more rows than columns and full column rank, then given noisy data b there is in general no solution, so the existence condition fails.

Even if the existence and uniqueness conditions are satisfied, which corresponds to A being invertible, then the stability condition may fail to hold. For the linear discrete inverse problem, this happens in case that A is *ill-conditioned*. The (2-norm) condition number of an invertible matrix is given by

$$\kappa(A) = \|A\|_2 \left\|A^{-1}\right\|_2, \tag{3.1}$$

and if $\kappa(A)$ is "small", i.e., not too far from the minimal value of 1, we say A is well-conditioned, whereas an A with "large" $\kappa(A)$ is called ill-conditioned. The condition number plays a central role for numerical stability. Assuming that the observed data, $\widetilde{b}$, is subject to an additive perturbation Δb, $\widetilde{b} = b + \Delta b$ of the ideal data, $b = Au$, we have the fundamental bound on the perturbation on the reconstructed image Δu, see, e.g. [62]:

$$\frac{\|\Delta u\|_2}{\|u\|_2} \leq \kappa(A) \cdot \frac{\|\Delta b\|_2}{\|b\|_2}. \tag{3.2}$$

For a small condition number, (3.2) ensures that small data perturbations can only lead to small reconstruction errors. But for a large condition number, say $\kappa(A) = 10^{10}$ or larger, even small data perturbations can lead to large reconstruction errors. In this case, it is unlikely that the naive solution $u = A^{-1}b$ will be even close to useful.

3.1.3 Regularization: Fixing the Hadamard conditions

One way to handle an ill-posed inverse problem is to introduce some further regularity to the problem in order to obtain a modified problem with a unique and stable solution. This approach is known as *regularization*. In case of an inconsistent linear system $Au = b$, the lack of existence of a solution can, for example, be fixed by replacing the problem by the least-squares problem

$$u^\star = \operatorname*{argmin}_u \|Au - b\|_2^2, \tag{3.3}$$

which has at least one solution for all A and b. If A has a nontrivial nullspace, this is, however, not enough to obtain a unique solution. We can achieve that by including an additional *regularization term*, for example, by Tikhonov regularization, see e.g., [69, 117], which takes the form

$$u^\star = \operatorname*{argmin}_u \left\{ \|Au - b\|_2^2 + \lambda^2 \|Su\|_2^2 \right\}, \tag{3.4}$$

where λ is the regularization parameter and S is a matrix, for example the identity or a discrete approximation of a derivative operator chosen to introduce smoothness [70]. Assuming the nullspaces of A and S have a trivial intersection, this problem has a unique solution. At the same time, this formulation helps alleviate the stability problem by carefully balancing the influence of the data with the stabilizing effect of the regularization term by proper selection of the regularization parameter λ.

More generally, $T(u) = \|Au - b\|_2^2$ and $R(u) = \|Su\|_2^2$ are examples of a *data fidelity* term and a *regularizer*, respectively, that make up the *regularized problem*

$$u_\lambda^\star = \operatorname*{argmin}_u \left\{ T(u) + \lambda R(u) \right\}. \tag{3.5}$$

The data fidelity measures the deviation between the measured data and a forward projected image. It can either be chosen based on a physical model of the measurement process or more heuristically. The use of the 2-norm in both terms, as in Tikhonov regularization, is attractive for several reasons. From a mathematical viewpoint, it leads to a problem that can be analyzed completely using standard linear algebra tools such as the singular value decomposition [69, 70] to obtain a closed-form solution. From a computational viewpoint, fast algorithms exist for determining the solutions, see e.g., [69]. It is, however, not clear that the 2-norm data fidelity is the best noise model. We will show in subsection 3.1.4 that the 2-norm data fidelity implicitly specifies a measurement process subject to Gaussian white noise.

The job of the regularizer is to incorporate any information about the solution that is available prior to acquiring any data. The regularizer should be chosen such that desirable images in the specific application are encouraged through a low value of the regularizer and undesirable images are penalized by a large value.

The regularizer can for example be chosen to be a p-norm ($p \geq 1$) of the signal u itself or of some transform applied to it, for example a wavelet transform or a discrete approximation of a derivative operator. Depending on the choice, images of different appearance and smoothness will be promoted. Of particular interest for the thesis work is the case of $p = 1$, which tends to encourage sparsity in the solution, i.e., few nonzeros. This choice is discussed further in chapter 4. A regularizer that has demonstrated potential for CT image reconstruction is *total variation* (TV), which is described in section 3.2.

Given a data fidelity and a regularizer, we still need to specify how to balance the emphasis on each, which is done through the choice of regularization parameter. The solution to the regularized problem depends strongly on the regularization parameter, and a natural question is what the best value is and how to find it.

Manually trying out a number of values followed by picking the best one seems unsatisfactory. Several more automated methods for selecting the regularization parameter have been suggested in the literature, including Morozov's discrepancy principle, use of the L-curve, generalized cross-validation, and use of the normalized cumulative periodogram, see [70] for an overview. These methods all make some assumptions, for example of the type of noise, and situations can occur where none of the methods reliably provide a good choice of regularization parameter. In such cases, it may be possible to use one's own knowledge of good parameter values obtained from previous similar cases, or it may be necessary to resort to manually searching for a good value.

It should be noted that a solution to the regularized problem is generally biased with respect to a solution to the unregularized discrete inverse problem. Even if we have a well-posed discrete linear inverse problem, i.e., a (stably) invertible A and ideal data b, then the regularized solution (for $\lambda > 0$) will *not* be equal to $A^{-1}b$. The size of the bias is governed by the regularization parameter. We can think of the bias as the price we must pay for making an ill-posed problem well-posed, so that, at least, an approximate solution of the original problem can be determined.

3.1.4 The Bayesian statistical perspective

A large branch of inverse problems takes a statistical approach, also known as a Bayesian approach, see, e.g., [21, 116]. In CT, this approach is sometimes also referred to as statistical image reconstruction, see, e.g., [58]. This framework describes signals statistically in terms of probability distribution functions (PDFs). Both the signal u and the observed data b are considered stochastic variables. The *likelihood function* $\pi_{\text{data}}(b|u)$ is a PDF that describes the likelihood (probability) of observing a specific outcome of b given knowledge of the signal u. The *prior* $\pi_{\text{prior}}(u)$ is another PDF that describes the kind of signals we are looking at. The goal is still to solve the inverse problem, i.e., determine what signal caused a given observation, which in the Bayesian formulation is described through a third PDF, the *posterior*, $\pi_{\text{post}}(u|b)$. The three PDFs are connected through Bayes' formula

$$\pi_{\text{post}}(u|b) \propto \pi_{\text{prior}}(u) \cdot \pi_{\text{data}}(b|u), \tag{3.6}$$

where "$\propto$" means "proportional to", since a normalization factor is left out.

To see the connection with the regularization approach, we assume that the data is subject to additive Gaussian white noise

$$b = Au + e, \tag{3.7}$$

where e is a noise vector with independent identically distributed (i.i.d.) elements from a normal distribution of zero mean and variance σ_e^2. Then e has the multivariate Gaussian distribution with PDF

$$\pi_{\text{data}}(e) = \frac{1}{(2\pi\sigma_e^2)^{N/2}} \exp\left(-\frac{1}{2\sigma_e^2}\|e\|_2^2\right). \tag{3.8}$$

Since $e = b - Au$ and e and u are assumed independent we can write

$$\pi_{\text{data}}(e) = \pi_{\text{data}}(e|u) = \pi_{\text{data}}(b - Au|u) = \pi_{\text{data}}(b|u), \tag{3.9}$$

and we have the likelihood function.

For the prior, we also assume that u has i.i.d. elements from a multivariate Gaussian distribution with zero mean and variance σ_u^2 and hence the PDF

$$\pi_{\text{prior}}(u) = \frac{1}{(2\pi\sigma_u^2)^{N/2}} \exp\left(-\frac{1}{2\sigma_u^2}\|u\|_2^2\right). \tag{3.10}$$

Given the likelihood function and the prior distribution we can compute the posterior distribution from Bayes' formula (3.6):

$$\begin{aligned}\pi_{\text{post}}(u|b) &\propto \frac{1}{(2\pi)^N(\sigma_u^2\sigma_e^2)^{N/2}} \exp\left(-\frac{1}{2\sigma_u^2}\|u\|_2^2 - \frac{1}{2\sigma_e^2}\|Au - b\|_2^2\right) \\ &\propto \exp\left(-\frac{1}{2\sigma_u^2}\|u\|_2^2 - \frac{1}{2\sigma_e^2}\|Au - b\|_2^2\right).\end{aligned} \tag{3.11}$$

From a Bayesian perspective, the posterior distribution itself is the solution to the inverse problem. One way to visualize the solution is to generate samples from the posterior distribution. Another option is to compute a point estimator of the solution and a standard choice is the maximum a posteriori (MAP) solution, which is the u with the highest posterior probability. We compute the MAP solution by maximization of the posterior:

$$u_{\text{MAP}} = \operatorname*{argmax}_u \ \pi_{\text{post}}(u|b) \tag{3.12a}$$

$$= \operatorname*{argmax}_u \ \exp\left(-\frac{1}{2\sigma_u^2}\|u\|_2^2 - \frac{1}{2\sigma_e^2}\|Au - b\|_2^2\right) \tag{3.12b}$$

$$= \operatorname*{argmax}_u \left\{-\frac{1}{2\sigma_u^2}\|u\|_2^2 - \frac{1}{2\sigma_e^2}\|Au - b\|_2^2\right\} \tag{3.12c}$$

$$= \operatorname*{argmin}_u \left\{\frac{\sigma_e^2}{\sigma_u^2}\|u\|_2^2 + \|Au - b\|_2^2\right\}, \tag{3.12d}$$

where to get to (3.12c) we take the logarithm, which does not change the maximizer, and to get to (3.12d) we multiply by $-2\sigma_e^2$, which replaces maximization

by minimization due to the negative sign. For $\lambda = \sigma_e/\sigma_u$ we have arrived at the simple Tikhonov problem (3.4) with S equal to the identity. That is, regularization can be interpreted as MAP-estimation in the Bayesian formulation, and this holds more generally than for Tikhonov regularization. Furthermore, we have now seen that simple Tikhonov regularization implicitly assumes Gaussian i.i.d. elements in both the image and the noise.

Gaussian white noise is not realistic for CT data. In a more realistic Gaussian noise model, the variance at each detector element can be matched to the variance of the logarithm-transformed Poisson distributed projection data, resulting in a weighted quadratic data fidelity. A different choice of data fidelity is the Kullback-Leibler (KL) divergence, aimed at transmission data,

$$T(u) = \sum_{i=1}^{M} \left[(Au)_i - b_i + b_i \log b_i - b_i \log(Au)_i \right], \tag{3.13}$$

which can be shown in the Bayesian framework to correspond to a multivariate Poisson distribution [4].

3.2 Total variation regularization

In this section we describe one specific kind of regularization, *total variation (TV)*, that has been the focus of much of the thesis work. We give a brief literature review and describe the background and motivation for the use of TV regularization in image processing and and discuss its desirable properties as well as some known drawbacks.

3.2.1 Definition in the continuous and discrete domains

TV was originally introduced in [105] for image denoising in a continuous-domain formulation. For a function $f(x_1, x_2)$ representing a 2D image over a domain Ω, the continuous (isotropic) total variation is given by

$$J_{\mathrm{TV}} f = \int_{\Omega} \sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2}\, dx_1 dx_2. \tag{3.14}$$

Assume that a clean signal f_{clean} is to be reconstructed from a noisy version $f_0 = f_{\mathrm{clean}} + \zeta$, where ζ is a white noise signal of zero mean and known standard

deviation σ. The denoising approach suggested in [105] consists of minimizing $J_{\mathrm{TV}} f$ subject to the constraints that

$$\int_\Omega f dx_1 dx_2 = \int_\Omega f_0 dx_1 dx_2 \quad \text{and} \quad \int_\Omega \frac{1}{2}(f - f_0)^2 dx_1 dx_2 = \sigma^2. \tag{3.15}$$

This problem is often referred to as the ROF-problem after its proposers Rudin, Osher and Fatemi.

This definition assumes that the image function is differentiable, but also non-differentiable functions are of high interest in imaging. For example, edges between different image regions are discontinuities and hence nondifferentiable. The definition of the continuous formulation TV can be extended to allow discontinuities as described, e.g., in [34, 122] and in the 1D case in [86]. We do not cover this extension here since our interest is the TV of a discrete image, for which there is no discontinuity problem.

The discrete definition of TV is obtained by replacing derivative operators by finite-difference approximations, for example forward differences, and the integral by summation. We can write the finite-difference approximations of derivatives in matrix-notation. For a 1D signal the discrete TV is defined as

$$R_{\mathrm{TV}}(u) = \|Du\|_1 = \sum_{j=1}^{N} |D_j u|, \tag{3.16}$$

where D_j is the finite-difference approximation of the derivative at the jth point and D is the matrix representing the combination of D_j for all values of j. When generalizing to higher dimensions, two different variants are possible, the *isotropic TV*,

$$R_{\mathrm{ITV}}(u) = \sum_{j=1}^{N} \|D_j u\|_2 , \tag{3.17}$$

corresponding to the continuous-domain formula (3.14), and the *anisotropic TV*,

$$R_{\mathrm{ATV}}(u) = \|Du\|_1 = \sum_{j=1}^{N} \|D_j u\|_1 , \tag{3.18}$$

for which a similar corresponding continuous-domain formula exists. D_j is now a finite-difference approximation of the spatial image gradient.

The problem considered in most of the thesis work is the discrete isotropic TV-regularized least-squares problem

$$u^\star = \underset{u}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Au - b\|_2^2 + \lambda R_{\mathrm{ITV}}(u) \right\}, \tag{3.19}$$

often without explicitly stating that the isotropic version is used.

3.2.2 Applications in image processing

TV-regularization has been studied extensively in the classical image processing disciplines of denoising, deblurring and inpainting, see, e.g., [34], [39].

The motivation for the interest in TV-regularization is the potential to obtain sharp edges, which is notoriously difficult for Tikhonov regularization (3.4). To see how TV acts differently on an edge than Tikhonov regularization, we consider the following example from [70] using a 1D continuous formulation. Consider the function

$$f_h(x) = \begin{cases} 0, & 0 \leq x < \frac{1}{2}(1-h), \\ \frac{t}{h} - \frac{1-h}{2h}, & \frac{1}{2}(1-h) \leq x \leq \frac{1}{2}(1+h), \\ 1, & \frac{1}{2}(1+h) < x \leq 1, \end{cases} \tag{3.20}$$

which is illustrated in Figure 3.1. The function has a linear transition centered around $h \in (0,1)$ and the slope is larger with smaller h. We compare how this

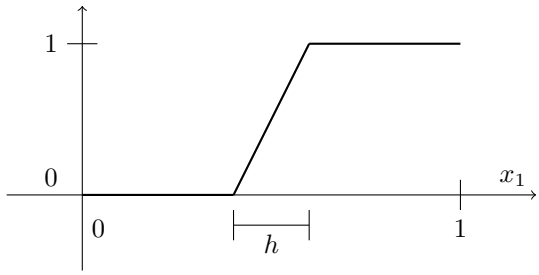


Figure 3.1: The function $f_h(x)$ from (3.20).

function is measured by the continuous TV-functional, and two functionals from Tikhonov regularization (3.4), namely the Euclidean norm of image function and of the derivative of the image function, corresponding to taking S to the identity and to the derivate operator in (3.4), respectively:

$$\|f_h'\|_1 = \int_0^1 |f_h'(x)|dx = 1, \tag{3.21a}$$

$$\|f_h'\|_2^2 = \int_0^1 f_h'(x)^2 dx = \frac{1}{h}, \tag{3.21b}$$

$$\|f_h\|_2^2 = \int_0^1 f_h(x)^2 dx = \frac{1}{2} - \frac{1}{6}h. \tag{3.21c}$$

The first of the three, the TV in (3.21a), is completely independent of h and hence the slope; it simply measures the magnitude of the jump. We see that this

is a property of the 1-norm, since when using the 2-norm in (3.21b) instead we get $1/h$, which increases rapidly toward infinity when h becomes smaller. If we omit the derivative, we get the result in (3.21c), which shows some dependence of h but as $h \to 0$ the value approaches the constant $1/2$, so a steep jump is not penalized nearly as hard as in the second case.

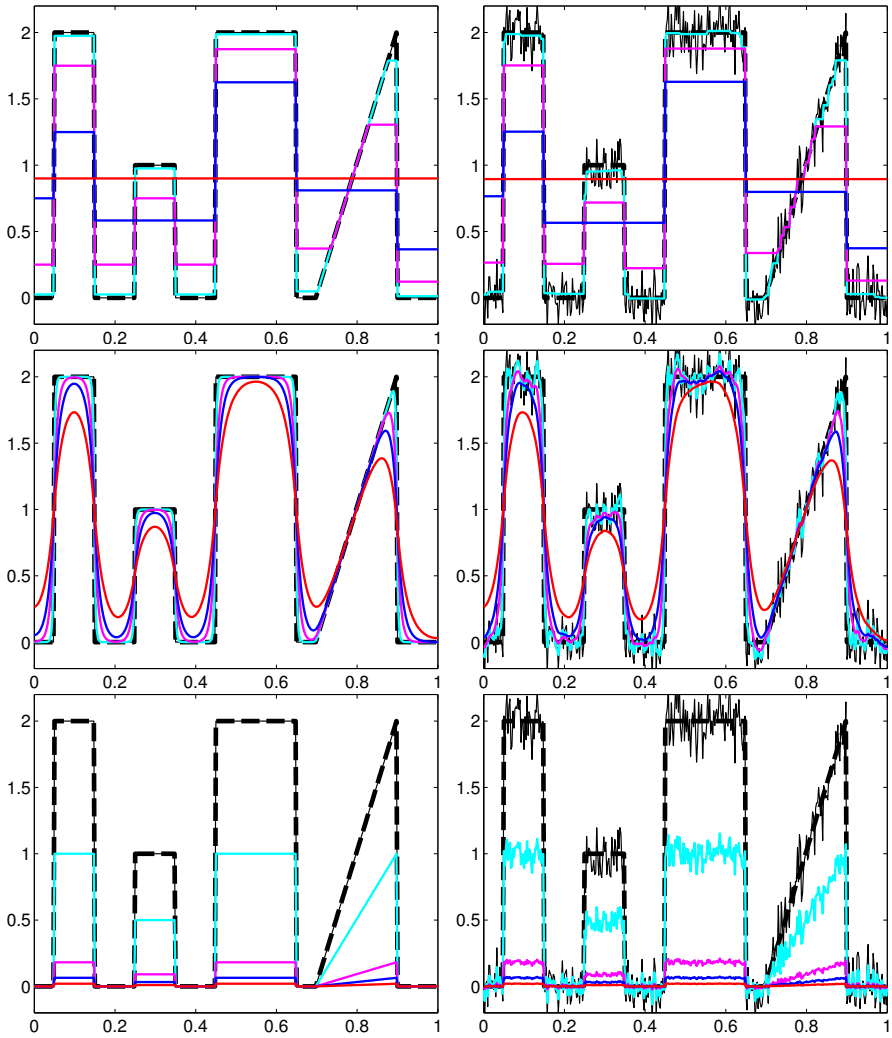


Figure 3.2: Top, middle, bottom: TV, Tikhonov-derivative-regularization, simple Tikhonov regularization. Left: denoised 1D signals from noise-free data. Right: Same, with 10% Gaussian noise. Thick, dashed = noise-free signal; thin = noisy signal; cyan,magenta,blue,red: denoised signals with $\lambda = [1, 10, 30, 100]$.

To see how these differences show up in practice, and to illustrate a few basic

properties of TV-regularized solutions, we consider a denoising example in the discrete setting, see Figure 3.2. A noise-free 1D discrete signal of 401 points on the interval $[0, 1]$ consisting mainly of piecewise constant parts and a linearly increasing part. The signal is subject to additive Gaussian noise and from the noisy observations the goal is to recover the original noise-free signal approximately. We compare the discrete versions of the three regularization methods. Two relative noise levels, 0% and 10%, are used; the former to illustrate performance in the theoretical noise-free scenario. The same four choices of the regularization parameter are used in all cases. Although the same values are not directly comparable among different regularizers we can see the general trends. We make the following observations:

- In the noise-free case, the TV-regularized solution of the piecewise constant regions are piecewise constant. The intensity depends on λ ranging from equal to the noise-free signal only at $\lambda = 0$ and more and more reduced as λ increases until finally only the constant image (with intensity equal to the mean value of the signal) is left. How much the intensity is reduced for each object appears to depend on the object width and not the original object intensity, since the first two narrow objects (of different original intensity) show the same absolute intensity reduction, while the third and twice as wide object shows only half the intensity reduction. These observations can be proven to hold, see [113].
- The Tikhonov-derivative-regularized solution is much smoother, which agrees with our previous observation that a steep transition is penalized hard by Tikhonov regularization of the derivative. The intensity also decreases with increasing λ.
- The simple Tikhonov-regularized solution manages to yield a piecewise constant solution, agreeing with a smaller h leading to a constant penalty. The intensity also decays with increasing λ.
- The noisy-data TV-regularized solutions on the piecewise constant regions are still piecewise constant, except for the case of the smallest considered λ. Even though not perfectly piecewise linear, the smallest choice of λ produces a fairly good approximation to the original signal. The linear part, which in the noise-free case is accurately recovered, now exhibits what is known as *staircasing artifacts*. As this example shows, staircasing is caused by the noise and is seen on parts of the object that are not piecewise constant.
- The noisy-data Tikhonov-derivative-regularized solutions show reduced noise but are also much smoother than the original signal.
- The noisy-data simple Tikhonov regularized solution shows no noise suppression, only the intensity reduction also present in the noise-free case.

It is precisely the property of allowing a piecewise constant solution with steep jumps while suppressing noise that has made TV successful in many image

processing applications. A steep jump in 1D corresponds to a sharp edge in two and higher dimensions. However, the staircasing artifacts (which for 2D images manifest in a "patchy" appearance of piecewise constant regions) and the intensity reduction (or contrast loss) can be unacceptably pronounced in some applications. As a result, it cannot be known in advance that TV-regularization is well-suited for a particular application simply because the solution is piecewise constant. In section 3.3 we address the application of TV-regularization to CT image reconstruction.

3.2.3 Computing the TV solution

The TV-solution is determined by solving an optimization problem, either the regularized version (3.19) or one of the constrained variants described in subsection 3.4.1. Numerous algorithmic strategies for solving optimization problems involving TV have been developed over the years, of which we mention a selection here to illustrate the variety of approaches that can be taken: The methods include time-marching schemes such as in the original TV-paper [105], fixed-point iteration [123], interior-point methods for second order cone programs [60], first-order methods [3, 7, 9, 39, 126] as well as our paper F, alternating minimization [124], duality-based methods [29, 31, 32, 127], split-Bregman [61], subgradient methods [2, 38], thresholding methods [7, 8, 41], domain-decomposition methods [59], graph-cut methods [30, 40] and other methods [47, 82].

Having a fast algorithm is important for obtaining an accurate solution in acceptable time. There is no simple answer to which algorithm is faster in general, because this question has many facets. For example, it may be that one algorithm produces an approximate solution faster than a second algorithm, but the second algorithm is faster in determining an accurate solution. Which algorithm to use depends on the accuracy needed in the particular application.

3.2.4 Generalizations

In an attempt to make use of the desirable properties of TV and reduce the artifacts such as staircasing and intensity reduction, numerous variants and generalizations have been proposed, including the use of a spatially varying and adaptively updated TV regularization parameter [113], "color TV" for denoising of vector-valued/color images [13], total generalized variation (TGV) using higher order derivatives [15], and TV-regularized 1-norm minimization for denoising with interestingly different geometric behavior [33].

3.3 Application to CT

3.3.1 Motivation: Low-dose imaging

In recent years, substantial attention has been given to the risk of radiation-induced cancer caused by CT scans, see e.g., [16, 42, 111]. It is believed that even a single CT scan increases the risk of developing cancer. In many cases, patients are subjected to a series of scans during initial diagnosing, treatment and follow-up examination, which considerably increases radiation exposure. Of particular concern are children, whose smaller bodies are more sensitive to radiation and for whom cancer has many living years to develop. Low-dose CT can help reduce the risk of radiation-induced cancer.

Another case where radiation dose is a concern is in the potential use of CT in screening programs, for instance for breast cancer. Breast cancer screening is commonly done by mammography—a procedure with a number of drawbacks including patient discomfort, limited sensitivity for cancer detection for example in dense breasts and inherent imaging difficulties such as tissue superposition in the conventional 2D display of 3D objects [83]. As an alternative, a dedicated breast CT scanner is under development [83]. It is clear that an effective screening program must detect more early-stage cancers than it induces itself in patients. This means challenging CT operating conditions as the total x-ray exposure is constrained to the order of mammography, which is much lower than for a conventional CT scan.

These concerns, among others, motivate the interest in low-dose CT imaging. However, as explained in subsection 2.1.2, a reduction in dose leads to reduced data quality. A standard clinical CT-scanner uses a rather high x-ray exposure [16] in order for its analytical inversion method to produce a high-quality reconstruction. As argued in [86], there is not much to be gained in terms of improved reconstruction quality by using other methods such as TV-regularization in a full-data case. But once the exposure is lowered, the analytical inversion methods start to produce undesirable artifacts, which motivates alternative reconstruction methods.

Exposure can be reduced either by reducing the x-ray intensity in each projection or by reducing the total number of projections. The latter approach amounts to reducing the number of rows in the CT system matrix, thereby obtaining an increasingly underdetermined linear system. This is the scenario considered by sparse image reconstruction methods, and we therefore focus on CT image reconstruction from a reduced number of projections.

3.3.2 TV-regularization in CT image reconstruction

As described in subsection 3.2.2, TV-regularization tends to work well on piecewise constant 1D signals, and in higher dimensions this corresponds to images with regions of constant intensity, also referred to as "blocky images" for example in [123]. The human body consists of fairly well-separated regions of similar tissue, which means that an image of the x-ray attenuation coefficient over a cross section of the body tends to be "blocky". For this reason it is natural to consider TV-regularization for CT image reconstruction.

TV-based image reconstruction and edge-preserving regularization were introduced to the CT community by various authors, e.g., [46, 54, 95, 108, 109]. These initial works spurred a multitude of works on applying TV-based methods to various specific CT applications, e.g., [12, 36, 67, 68, 81, 104, 110, 112] and development of optimization algorithms suited for the large scale of practical CT, e.g., [45, 101, 102] and our paper E.

There are many examples where TV-regularized image reconstruction leads to a significant reduction of radiation exposure, while maintaining or even improving reconstruction quality compared to analytical reconstruction. One example is [12], from which a selection of reconstructions is shown in Figure 3.3 (visually adjusted, courtesy of X. Pan, University of Chicago). In that work, the authors consider reconstruction of a 3D physical test phantom from cone-beam data. They demonstrate that a TV-reconstruction provides a better reconstruction of small low-contrast disk-shaped objects from 10 times fewer projections than the standard analytical FDK algorithm. In comparison, the FDK reconstruction from the same number of projections is heavily corrupted by streak artifacts. Examples such as this one show great promise for a considerable dose-reduction through the use of TV-based image reconstruction.

3.4 Optimization and algorithm considerations

The regularized problem (3.5) is an optimization problem. Other optimization problem formulations can also be of interest for regularization. In this section we present some general aspects of optimization and optimization algorithms relevant for regularization.

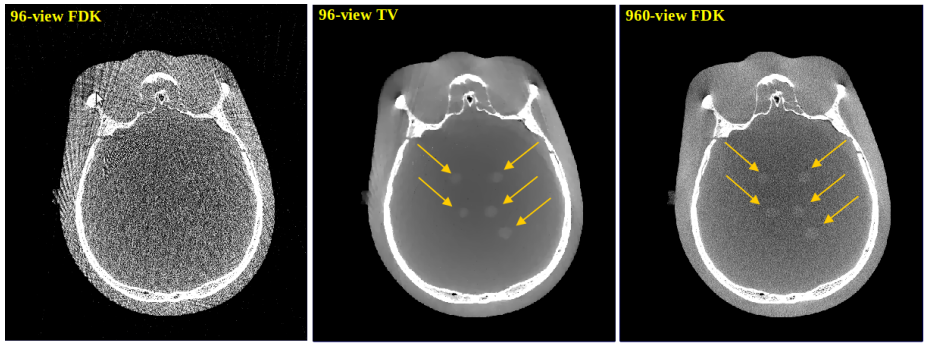


Figure 3.3: Slices through 3D reconstructions of a physical head phantom from a circular cone-beam CT scan. Adapted from [12], with permission from X. Pan, University of Chicago. Left: Reduced-data (96 projections) analytical FDK-reconstruction, center: reduced-data (96 projections) TV-reconstruction, right: full-data (960 projections) analytical FDK reconstruction. Arrows indicate low-contrast disk-shaped objects of interest.

3.4.1 Properties of optimization problems

In this subsection we give a general overview of properties that are useful to be aware of when formulating an optimization problem to provide regularization to a given imaging problem.

Unconstrained vs. constrained: Different variants of the optimization problem can be considered. We already mentioned the *regularized* or *penalized* version (3.5). Another possibility is the *data-constrained* or typically simply referred to as the *constrained* formulation

$$u_\epsilon^\star = \operatorname*{argmin}_u \; R(u) \tag{3.22a}$$

$$\text{s.t.} \quad T(u) \leq \epsilon, \tag{3.22b}$$

where ϵ is the constrained regularization parameter. This problem can be more intuitive to work with than the regularized version, because in the constrained version the regularizing effect is obtained by putting a restriction on the data fidelity term. If for example the noise level is known, then ϵ can be selected more directly based on this knowledge, whereas λ in the regularized formulation can not be chosen directly based on this. Also, in contrast to the regularized formulation, the data-constrained formulation allows us to study the the idealized case of an equality constraint on the data misfit by taking $\epsilon = 0$. For example, in our paper A we use the data-constrained formulation of TV-regularization,

e.g.,

$$u_\epsilon^\star = \underset{u}{\operatorname{argmin}} \; R_{\mathrm{TV}}(u) \tag{3.23a}$$

$$\text{s.t.} \quad \|Au - b\|_2 \leq \epsilon, \tag{3.23b}$$

to study the reconstruction quality as $\epsilon \to 0$.

Sometimes, but not as often, the roles of the terms are switched to obtain what we will call the *regularizer-constrained* formulation

$$u_\tau^\star = \underset{u}{\operatorname{argmin}} \; T(u) \tag{3.24a}$$

$$\text{s.t.} \quad R(u) \leq \tau, \tag{3.24b}$$

in which τ is the regularizer-constrained regularization parameter.

Interestingly, the three formulations, i.e., the regularized (3.5), data-constrained (3.22) and regularizer-constrained (3.24) problems, can be seen as equivalent in the sense that they yield the same solutions, however the connection between the parameters λ, ϵ and τ is not known in advance and depends in a nontrivial way on the data. In some cases it is possible to use one formulation to obtain a solution to a different formulation. Thanks to Pareto optimality [14, 120], if we know the optimal regularized solution $u_\lambda^\star$, we can easily compute the corresponding ϵ and τ for which the same image is the solution to each of the constrained formulations. This is the case when the constraint is satisfied with equality, i.e., we can compute $\epsilon = T(u_\lambda^\star)$ and $\tau = R(u_\lambda^\star)$. We are using this trick in our paper A to compute a data-constrained solution by use of an algorithm designed for the regularized problem.

For all three cases, it is possible to include further constraints on the image; we will simply write $u \in C$. A simple example is to enforce nonnegativity on the solution by taking $C = \mathbb{R}_+ \cup \{0\}$.

Convex vs. nonconvex: Rarely does an optimization problem have a closed-form solution and hence a numerical algorithm must be used to compute the solution. For the completely general class of optimization problem there is no guarantee that a solution can be computed numerically and we therefore restrict our attention to more narrowly defined classes. For example, nonconvex problems are notoriously hard to deal with because there can be many local minimizers. A local minimizer is what optimization algorithms aim at producing and from having obtained a local minimizer there is in general no way to know whether other local minimizers exist and whether the global minimum is attained. In contrast, convex problems enjoy the very useful property that any local minimizer is also a global minimizer.

It is however well-known that nonconvex problems in some cases lead to "superior" solutions compared to convex problems. One such example is considered in chapter 4: nonconvex "p-norms" with $p < 1$ are better at promoting sparsity in an image than the closest convex problem of taking $p = 1$. This can, for example, be used to construct a nonconvex variant of TV-regularization. If in the 1D case of (3.16), we replace the 1-norm with a "p-norm" with $p < 1$, i.e.,

$$R^p_{\mathrm{TV}}(u) = \|Du\|_p = \left(\sum_{j=1}^{N} |D_j u|^p \right)^{1/p}, \tag{3.25}$$

we obtain a regularizer that promotes sharp edges even better than TV, see, e.g., [35, 92] and our paper H for examples in the setting of CT. Due to the nonconvexity, there is however no guarantee that the numerical algorithm will not get stuck in an undesirable local minimizer.

In the thesis work we restrict ourselves to consider convex problems, i.e., the objective function is a convex function and constraints specify a convex set, with the single exception of our paper H.

Smooth vs. nonsmooth: We say that a function is smooth if it is continuously differentiable. The class of smooth optimization problems is well-established, especially in terms of algorithms. The well-known gradient or steepest descent method is the basic choice, while for twice continuously differentiable problems, methods such as Newton's method using the Hessian are typically faster for up to moderately-sized problems.

Many optimization problems of interest to CT, however, are nonsmooth, for example the TV-regularized least-squares problem in (3.19). The source of nonsmoothness here is the Euclidean norm, which is not differentiable at the origin. This problem is often circumvented by squaring the norm, but in the TV-functional this is not the case. A common strategy for approximately solving a nonsmooth optimization problem consists of working instead with a smoothed version and apply standard algorithms for smooth optimization. For the example of the TV-regularized least-squares problem, the Euclidean norm in the TV-function can be replaced for example by the Huber functional, see e.g. [122],

$$u^\star = \operatorname*{argmin}_u \left\{ \frac{1}{2} \|Au - b\|_2^2 + \lambda R^h_{\mathrm{HTV}}(u) \right\}, \tag{3.26}$$

where

$$R^h_{\mathrm{HTV}}(u) = \sum_{j=1}^{N} \Phi_h(D_j u) \tag{3.27}$$

and

$$\Phi_h(z) = \begin{cases} \|z\|_2 - \frac{1}{2}h, & \text{if } \|z\|_2 \geq h, \\ \frac{1}{2h}\|z\|_2^2, & \text{else.} \end{cases} \tag{3.28}$$

An advantage of this approach can in some cases be faster convergence, as discussed in subsection 3.4.2. On the other hand, the obtained solution is only an approximate solution of the original problem and the extra parameter introduced to control the "amount" of smoothing must be chosen in a way so that the problem becomes faster to solve while not distorting the solution too much.

The alternative to smoothing is to use algorithms specifically designed for nonsmooth optimization, which can be more challenging to work with, as well as being subject to slower convergence, but will on the other hand deliver the solution to the original nonsmooth problem.

In the thesis work we consider both the strategy of smoothing and of using nonsmooth optimization algorithms. Specifically, our paper F takes the approach of smoothing the TV-functional, while in the paper A we study an algorithm for nonsmooth optimization, for example applied to the TV-regularized least-squares problem.

Uniqueness vs. nonuniqueness of solution: We often speak of *the* solution to an optimization problem, but even for a convex problem the solution is not necessarily unique. One such example is the problem

$$u^\star = \operatorname*{argmin}_u \|u\|_1 \quad \text{s.t.} \quad Au = b, \tag{3.29}$$

which we consider in chapter 4. A sufficient, but not necessary, condition for existence of a unique solution is the stronger notion of strict convexity. We will consider both problems with a unique solution (paper A, paper C, paper E and paper F) and with a solution set of more than one point (paper B and paper C).

3.4.2 Optimization algorithms

In order to solve the reconstruction optimization problem we need to use a numerical algorithm. There is a huge selection of optimization algorithms available in the literature so it is not clear which algorithm is going to be the "best choice" for a given problem. Ideally, we want the algorithm to be fast, have low memory requirements and produce an accurate solution. No single algorithm is the "best choice" in all cases. We discuss a few aspects to be aware of.

First of all, we ask that an algorithm is provably convergent to the minimizer. If this is not the case, then there is no reason to expect that the algorithm will solve the optimization problem. In practice, however, there are examples of algorithms that have not been proven convergent but are nevertheless successful in computing the solution in some cases, for example the ASD-POCS algorithm [109].

During the development phase, where a given optimization problem is being considered for at specific purpose, it is useful to study properties of the solution as function of the problem parameters, knowing that the solution really is the solution and not polluted by inaccuracies. For this purpose it can be relevant to use general-purpose software such as CVX [64, 75] for convex problems and MOSEK [85] for linear, quadratic and conic programs. Their use is restricted to fairly small problems, but the computed solutions are very reliable, i.e., not very sensitive to rounding errors and other numerical difficulties that can be a challenge for more specialized and less mature software. Both CVX and MOSEK are based on interior-point methods, and the limiting factor for the problem size is the need to factorize a matrix at least of size equal to the number of pixels squared. For example, a 1024^2-pixel 2D image, this leads to a matrix of size $\approx 10^6 \times 10^6$, which in double precision would require around 8 terabytes memory, clearly out of the question for most computer systems.

For the problem sizes that are feasible to handle, the advantages of general-purpose algorithms are the reliability in computing an accurate solution, the flexibility to solve a variety of optimization problems of interest and the reports about the quality of the computed solution. Another use of general purpose software is to compute a reliable reference for verification of other optimization algorithms.

For obtaining a solution to a realistically sized reconstruction problem in acceptable time we cannot use general-purpose software and consider instead more specialized algorithms.

For the sizes of the optimization problems arising in CT, we cannot use second-order methods, such as Newton's method and variants, for the same reason that interior-point methods are impractical. That leaves us with only using first-order information. Furthermore, the size of the system matrix in realistically-sized problems makes it infeasible to store in memory. We will only be able to compute the result of applying A or its transpose to a vector, so the choice of optimization algorithm is restricted to only algorithms involving matrix-vector products. Even with these restrictions, there are many different possible algorithms to choose from. We do not intend to give a complete overview and comparison of algorithms, rather we simply describe a few selected algorithms, which have been subject to study in the thesis.

The simplest first-order method is the gradient method for smooth, unconstrained problems, also known as the steepest-descent method. For the optimization problem

$$u^\star = \operatorname*{argmin}_u F(u), \tag{3.30}$$

the gradient method simply iterates

$$u^{(k+1)} = u^{(k)} - h^{(k)} \nabla F(u^{(k)}), \tag{3.31}$$

where the superscript (k) indicates the iteration number. The gradient method has a natural extension to constrained problems through the use of a projection operator: If C is the set of constraints and P_C the Euclidean projection onto C, then the gradient projection method consists of the iteration

$$u^{(k+1)} = P_C \left(u^{(k)} - h^{(k)} \nabla F(u^{(k)}) \right). \tag{3.32}$$

The step length $h^{(k)}$ can either be constant or chosen adaptively in each iteration. The gradient method is simple and intuitive, simply take a step in the direction of the largest (negative) gradient and continue until the minimizer is reached. However, the method is known to be slow, typically too slow to be practical for CT problems: it has worst-case convergence rate $F^{(k)} - F^\star \leq O(1/k)$, see e.g., [14]. The big-O notation for worst-case convergence rate means a decay rate *at least* proportional to the given rate, here $1/k$, with an unknown constant, but possibly faster.

Many modifications of the basic gradient method are considered in the literature. One particular variation [103], which involves a step-length selection proposed by Barzilai and Borwein [6] to yield a scalar approximation to the Hessian of $F(u)$, and a nonmonotone line search [65], has been demonstrated to provide a significant speed-up on many problems, although it shares the same worst-case convergence rate as the basic gradient method. We study and implement a variation of this method in our paper F. In that paper we also consider and implement a so-called *accelerated* gradient method proposed by Nesterov [90, 91]. This method enjoys a faster worst-case convergence rate of $F^{(k)} - F^\star \leq O(1/k^2)$, which is possible due to the use of an auxiliary sequence of iterates.

As mentioned, the gradient method (and variants) are designed for a smooth problem, so in applying it to the TV-regularized least-squares problem, it is necessary to smooth the TV-functional. Another option is to keep the TV-functional nonsmooth and apply a generalization of the gradient method called the *subgradient method*, in which the gradient is simply replaced by a subgradient. However, the worst-case convergence rate of the subgradient method is $F^{(k)} - F^\star \leq O(1/\sqrt{k})$, [121].

Yet another option is a different generalization of the gradient method known as the *proximal gradient method*. The proximal gradient method works with an unconstrained problem with two terms, one smooth and one nonsmooth. The nonsmooth term is replaced by its so-called prox-operator. The proximal gradient method retains the worst-case convergence rate of $F^{(k)} - F^\star \leq O(1/k)$, but to be efficient it requires the prox-operator to be cheap to evaluate, which may not always be the case. The proximal gradient method can also be equipped with the acceleration technique of Nesterov to achieve the worst-case convergence rate of $F^{(k)} - F^\star \leq O(1/k^2)$, see [7].

Many other algorithms exist; a few important examples are the augmented Lagrangian method [11] and the alternating direction method of multipliers (ADMM) [51], which is also referred to as split-Bregman by some authors [61]. General overviews of first-order methods and worst-case convergences rates can be found in [90, 119].

3.4.3 Where to start and when to stop?

The optimization algorithms just described are all iterative procedures and must be initialized from some point as well as terminated at some point. For the initial point, if we assume the solution is unique, we note that as long as the algorithm is convergent, the choice of initial point does not affect the determined solution to the optimization problem, which will be achieved if the algorithm is run for long enough. However, if the same algorithm used on the same problem but with different initial points are terminated early of convergence, then the two computed solutions may be different. If the solution is not unique but consists of a set of points, then the particular solution selected from the solution set depends on the initial point.

The worst-case convergence rates are unaffected by the choice of initial point, but in practice an informed initial point (closer to the minimizer than an uninformed one, such as the zero vector) can reduce the number of iterations needed significantly. This idea can be exploited as a *warm-starting strategy*: If we have solved an optimization problem of the regularized type (3.5) with some choice of regularization parameter $\lambda = \lambda_0$, then, using this solution, we can likely obtain the solution for a slightly different $\lambda = \lambda_0 + \Delta\lambda$ a lot faster than if we must start from an uninformed initial point.

Optimization theory provides optimality conditions that must be satisfied by a minimizer and be used to check whether an iterate in an algorithm is close to the minimizer. The simplest case is for an unconstrained problem, where an optimality condition is $\nabla F(u) = 0$, which gives the termination criterion:

Terminate after iteration k if $\left\|\nabla F(u^{(k)})\right\|_2 < \eta$, for some user-specified tolerance parameter η on the required accuracy of the solution. For constrained optimization problems, various optimality conditions exist that can be used to construct termination criteria; some important examples are the Karush-Kuhn-Tucker (KKT) conditions and the duality gap of primal-dual methods, see e.g. [14]. However, it should be noted that in general it is hard to make any quantitative statements on how small the tolerance η should be set to obtain a solution of a given accuracy.

3.5 Summary

We have now given a brief introduction to the field of inverse problems and demonstrated how regularization can be employed to obtain a meaningful solutions in case of ill-posed or ill-conditioned inverse problems. We gave particular emphasis to total variation (TV) regularization, which is known to enable reconstruction of images with sharp edges. We demonstrated the regularizing effect of TV on a 1D denoising numerical example. The example illustrated the edge-preserving behavior of TV but also revealed some of the well-known drawbacks such as the staircasing effects and loss of intensity. We also presented the application of TV-regularization to low-dose CT image reconstruction. In the next chapter we introduce the perspective of sparse image reconstruction, which can shed some light on the use of, for example, TV-regularization for image reconstruction.

Chapter 4

Sparse image reconstruction

In the past decade there has been a large interest in reconstruction methods for images with a sparse representation, i.e., that have relatively few nonzero representation coefficients, or that are sparse under some transform applied to the image. The interest stems from the potential to recover the image or signal from fewer measurements than what is required for reconstructing a general signal. In this chapter we give an introduction to this large and rapidly growing field. The presentation is based on [18, 28, 52, 55].

4.1 Sparse solutions of linear systems

4.1.1 Selecting one solution among the many

Consider the generic discrete inverse problem of recovering a discrete image u from data b obtained through a process modeled by the measurement matrix $A \in \mathbb{R}^{M \times N}$:

$$Au = b. \tag{4.1}$$

This corresponds to the DD imaging model from section 2.2. If A is square and invertible, then $u = A^{-1}b$ is the unique solution. If $M > N$ and A has full

column rank, then the solution is also unique and given by

$$u_{\mathrm{LS}} = (A^T A)^{-1} A^T b. \tag{4.2}$$

The subscript LS stands for "least-squares", because in case of an inconsistent linear system, such that $Au = b$ does not have a solution (for example for a noisy b), then u_{LS} is the unique signal that minimizes the squared residual 2-norm $\|Au - b\|_2^2$.

The final case, and the one we focus on here, is when $M < N$, where infinitely many solutions exist due to A having a nontrivial nullspace. One way to specify a unique solution is by taking the so-called minimum-norm solution, i.e.,

$$(P_2) \qquad u_{P_2} = \underset{u}{\operatorname{argmin}} \|u\|_2^2 \quad \text{s.t.} \quad Au = b, \tag{4.3}$$

which in case of A having full row rank is uniquely given by

$$u_{P_2} = A^T (AA^T)^{-1} b. \tag{4.4}$$

This choice of obtaining a unique solution is convenient, since it has a closed-form expression and in addition can be analyzed completely through the use of standard linear algebra tools such as the singular value decomposition. More generally, without any restrictions on the rank of A, we note that the solutions u_{P_2} and u_{LS} can both be expressed as $A^\dagger b$, where $A^\dagger$ is the Moore-Penrose pseudo-inverse of the system matrix A.

A different choice is to ask for the most *sparse* solution, i.e., the solution to $Au = b$ having the smallest number of nonzero elements. The sparsity of a signal u is typically measured in the "0-norm", $\|u\|_0$, which simply counts the number of nonzero elements. A signal with s or fewer nonzeros is called s-sparse. Referring to $\|\cdot\|_0$ as the 0-norm is misleading, since it does not satisfy the positive scalability property required of a norm, but we use this notation to be consistent with the literature. The most sparse solution can be written

$$(P_0) \qquad u_{P_0} = \underset{u}{\operatorname{argmin}} \|u\|_0 \quad \text{s.t.} \quad Au = b. \tag{4.5}$$

As will be demonstrated in what follows, the (P_0) and related problems lead to useful solutions when the signal is known to be sparse.

Unfortunately, the task of determining the solution to (P_0) is by no means easy. In fact, it is a problem of combinatorial complexity, because it calls for trying out all combinations of k-sparse vectors, starting from $k = 1$ and continuing to increase k until a solution is found. Hence, in general, for problems of practical interest we cannot determine the solution. There are two general strategies for attempting to determine the solution: greedy methods and relaxation into a convex optimization problem, both of which we describe briefly here.

4.1.2 Greedy methods

The basic greedy method, known as orthogonal matching pursuit (OMP), see, e.g., [118], for determining the most sparse solution to $Au = b$ begins from the zero signal u and an empty set of indices of nonzero (support) elements. The following steps are then iterated: A search is done for the single element of u that yields the smallest $\|Au - b\|_2$ while keeping remaining elements of u fixed at zero or their value determined in previous iterations. The found element's index is included in the support set, and support elements of u are updated to the values that minimize $\|Au - b\|_2$, while keeping nonsupport elements fixed at zero. The contribution from the updated u to the data is subtracted out of the data, before proceeding to the next iteration. The process continues to the remaining data residual norm is below a user-specified termination threshold or a maximal number of iterations is reached. With each iteration a single new nonzero element is introduced in u, so after s iterations, an s-sparse signal is found.

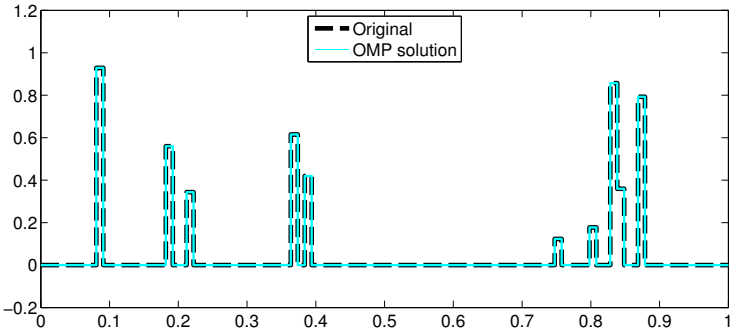


Figure 4.1: Orthogonal matching pursuit (OMP) can in some cases determine a sparse solution of (P_0). In the shown example, a 10-sparse signal of length 100 is recovered from 50 measurements.

An example of using OMP to compute a sparse solution is given in Figure 4.1, where a 10-sparse vector of length 100 is to be recovered from measurements obtained through a 50×100 matrix with elements drawn from the uniform distribution over $[0, 1]$. The OMP solution (full cyan) recovers the original accurately.

The greedy strategy of selecting in each iteration the best new element to include leads to a much faster procedure than the exhaustive combinatorial search required in general to solve (P_0). A major drawback is that the greedy strategy can, and will in many cases, fail to recover the original signal by selecting to include an element, which is not in the support of the original signal. Once an element is in the support, it can not be expelled again, and therefore this can

lead to a result far from the most sparse solution. Many variants of this basic greedy method exist with potentially better performance than the simple OMP, but all come with the same fundamental risk that the greedy strategy may fail.

4.1.3 Relaxation methods

The second strategy, which is known as relaxation, consists of replacing the 0-norm by an approximation that is easier to work with. The most commonly used approximation is the 1-norm, which leads to replacing (P_0) by

$$(P_1) \qquad u_{P_1} = \underset{u}{\operatorname{argmin}} \|u\|_1 \quad \text{s.t.} \quad Au = b. \tag{4.6}$$

For $p \geq 0$ the p-norm is defined as

$$\|x\|_p = \left(\sum_j |u_j|^p \right)^{1/p}, \tag{4.7}$$

and for values of $p \in [0, 1)$, we will still refer to $\|x\|_p$ as the p-norm, even though just like the 0-norm, it is not a norm. In Figure 4.2 we show 1D versions of a selection of p-norms. It is the behavior of a function on argument values close to 0 that determines whether the function is sparsity-promoting. The 2-norm is very flat around 0, so small but nonzero values give a relatively small contribution to the 2-norm of the signal. The kink of the 1-norm means that small values have a relatively larger contribution to the norm of the signal. The smaller p becomes, the larger is the relative contribution by small elements compared to that of elements with larger absolute value. This leads to a sparsity-promoting effect which is more pronounced the smaller p is. The limiting case is $p = 0$, where any nonzero element has the same contribution of 1 and only a zero-element contributes by 0 to the value of the 0-norm. So if we want to promote sparsity

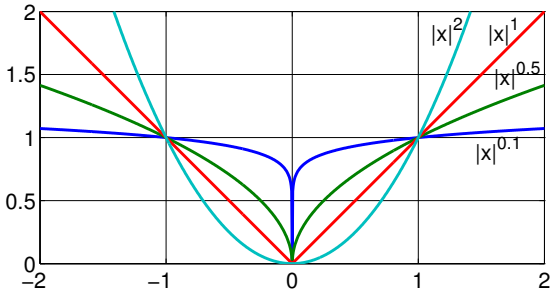


Figure 4.2: Comparison of scalar version of p-norms for $p = 2, 1, 0.5, 0.1$.

we should ideally use the smallest p possible. The drawback of $p < 1$, however, is that the "norms" are not convex, which as described in chapter 3, leads to much harder optimization problems. The smallest p for which we get a convex optimization problem is $p = 1$. Since the 1-norm has some sparsity-promoting behavior due to its kink and is convex, it is a commonly used trade-off between sparsity-promotion and computational tractability coming from convexity.

In most of our thesis work we have focused on convex optimization problems but it is widely recognized, and hopefully clear from the given example, that nonconvex functions can be more sparsity-promoting. In our paper H we study a nonconvex optimization problem and do find better sparsity-promoting behavior than the corresponding convex problem. It comes, however, at the price of potentially introducing local minima and the problem of determining whether a numerically computed solution is indeed the desired global minimizer or a sub-optimal, and not maximally sparse, local minimizer.

To illustrate geometrically why (P_1) produces a sparse solution, while (P_2) does not, we consider a tiny example of finding a 2-element vector from a single measurement through the measurement matrix $A = [1, 2]$ and the datum $b = 2$. The (P_1) and (P_2) solutions are illustrated in Figure 4.3 along with the 1-norm and 2-norm disks. The blue line represents the solution space of $A(x, y)^T = b$ and the (P_1) and (P_2) solutions can be found geometrically by inflating the respective norm disks from the origin until the disk touches the line. For (P_1), due to the kink of the 1-norm, this happens on the y-axis leading to a sparse solution $(0, 1)$. For the isotropic 2-norm disk, the solution is not on a coordinate axis and hence not sparse.

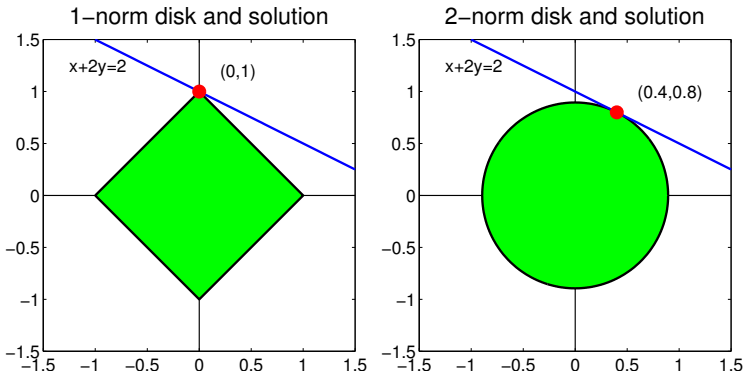


Figure 4.3: The problem (P_1) tends to produce a sparse solution, while (P_2) tends to produce a nonsparse solution.

As a final example, we consider reconstruction of the same sparse signal as

in Figure 4.1 through (P_1) and (P_2). The resulting solutions are shown in Figure 4.4. As expected the (P_2) solution is not sparse, even if it manages to produce a decent approximation of the nonzero values. The (P_1) solution on the other hand is sparse and recovers in fact, up to numerical accuracy, the original signal.

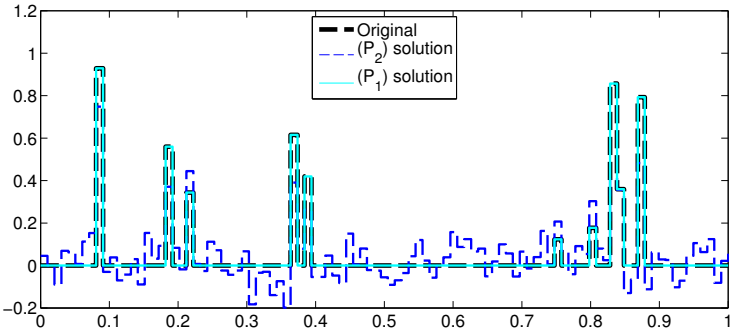


Figure 4.4: Comparison of (P_1) and (P_2) solutions. The (P_1) solution is sparse and accurately reconstructs the original, while the (P_2) solution is nonsparse.

4.2 Extensions of the basic sparsity problem

4.2.1 Sparsity in other representations

If the signal u itself is not sparse but it has a sparse representation, for example in a wavelet basis, $u = \Phi c$, where the matrix Φ holds the basis elements in its columns, then we can modify the problem to minimize the 0-norm of the coefficients:

$$\text{(Synthesis)} \qquad u^\star = \underset{c}{\operatorname{argmin}} \|c\|_0 \quad \text{s.t.} \quad A\Phi c = b. \tag{4.8}$$

This problem is referred to as the *synthesis* formulation [52], because the signal is being synthesized from the coefficients c. If Φ is orthogonal, corresponding to an orthonormal basis, then $c = \Phi^T u$ and the *analysis* formulation

$$\text{(Analysis)} \qquad u^\star = \underset{u}{\operatorname{argmin}} \left\|\Phi^T u\right\|_0 \quad \text{s.t.} \quad Au = b, \tag{4.9}$$

is equivalent to the synthesis formulation. It is called "analysis" because in this case the signal is being analyzed into its coefficient by Φ^T. As before, for approximately solving either of these problems, the 0-norm is normally replaced by

the 1-norm. If Φ is not a square, invertible matrix, then the synthesis and analysis formulation are different. Further discussion on similarities and differences between the synthesis and analysis formulations are given in [52, 53, 87].

4.2.2 Sparsity after application of transform

A related but slightly different situation occurs if we have a signal u that is sparse after applying a transform $Q(\cdot)$ to it. The relevant optimization problem is then

$$u^\star = \underset{u}{\operatorname{argmin}} \|Q(u)\|_0 \quad \text{s.t.} \quad Au = b. \tag{4.10}$$

If Q is a linear operator, $Q(u) = Qu$, then by taking $\Phi^T = Q$ we have a special case of the analysis formulation (4.9). A relevant example occurs when taking Q to be the discrete forward difference approximation of the gradient D. Together with replacing the 0-norm by the 1-norm, this leads to the anisotropic total variation (TV) minimization problem (see subsection 3.2.1),

$$u_{\mathrm{ATV}} = \underset{u}{\operatorname{argmin}} \|Du\|_1 \quad \text{s.t.} \quad Au = b. \tag{4.11}$$

To obtain the isotropic TV problem, we need the nonlinear transform Q_{ITV} of computing the 2-norm magnitude of the discrete gradient at each pixel j,

$$[Q_{\mathrm{ITV}}(u)]_j = \|D_j u\|_2, \qquad j = 1, \ldots, N, \tag{4.12}$$

and the isotropic TV minimization problem becomes

$$u_{\mathrm{ITV}} = \underset{u}{\operatorname{argmin}} \|Q_{\mathrm{ITV}}(u)\|_1 \quad \text{s.t.} \quad Au = b. \tag{4.13}$$

Comparing with (3.19) and (3.17) we see that

$$R_{\mathrm{TV}}(u) = \|Q_{\mathrm{ITV}}(u)\|_1 = \left\| \begin{pmatrix} \|D_1 u\|_2 \\ \vdots \\ \|D_N u\|_2 \end{pmatrix} \right\|_1, \tag{4.14}$$

so isotropic TV is the 1-norm of the 2-norm gradient magnitudes.

4.2.3 Relaxing the equality constraint

So far, we have only presented problems involving the strict equality constraint $Au = b$. The equality-constrained problem is of high theoretical interest but

represents an idealized problem. From a practical perspective, it is clear that the measured data is never ideal and consistent with the model so some misfit must be allowed. This can be done by applying the regularized, data-constrained and regularizer-constrained formulations from subsection 3.4.1.

The problem (P_1) can, for example, be modified to allow some data inconsistency through the regularized formulation

$$u_{P_1^\lambda} = \underset{u}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Au - b\|_2^2 + \lambda \|u\|_1 \right\}, \tag{4.15}$$

where λ is the regularization parameter. This problem is also known as *basis pursuit denoising* (BPDN), see e.g. [37]. An alternative is the data-constrained formulation with parameter ϵ:

$$(P_1^\epsilon) \qquad u_{P_1^\epsilon} = \underset{u}{\operatorname{argmin}} \|u\|_1 \quad \text{s.t.} \quad \|Au - b\|_2 \leq \epsilon. \tag{4.16}$$

Another example occurs if we use the regularized formulation with a 2-norm data fidelity term, sparsity in the gradient magnitude and the 1-norm instead of the 0-norm. We obtain the TV-regularized least-squares problem

$$u^\star = \underset{u}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Au - b\|_2^2 + \lambda R_{\mathrm{TV}}(u) \right\}, \tag{4.17}$$

where the factor of $1/2$ is introduced for convenience in computing the gradient.

4.3 Theoretical recovery guarantees

So why are we so interested in finding sparse solutions to underdetermined linear systems? The answer is that if we know that the signal we are trying to uncover is sparse (or has a sparse representation or is sparse after applying a transform), then there is hope that we can do so using fewer measurements than for a nonsparse signal. In other words, we can reduce the sampling effort while still obtaining a good reconstruction, assuming the image is sparse. For this to be possible, there are certain conditions on the measuring process, in particular the measuring matrix A. This section describes some theoretical results establishing a connection between the sparsity and the sampling required for accurate reconstruction. There are two different but closely related perspectives: The sparse representation perspective and the compressed sensing (CS) perspective; we begin by describing the former.

4.3.1 Sparse representation perspective

A fundamental property of a matrix A is its *spark*, $\text{spark}(A)$, which is defined as the smallest number of columns that are linearly dependent. Note that, while it has some similarities with the rank of a matrix, i.e., the *largest* number of linearly *in*dependent columns, it is very different. A fundamental condition on uniqueness of the most sparse solution can be expressed using the spark: If an s-sparse solution $u^\star$ exists to $Au = b$ and $\text{spark}(A) < 2s$, then $u^\star$ is the unique most sparse solution possible. Thus, if we find a solution u to $Au = b$, by whatever method, and it is s-sparse and we happen to know that the spark of A is smaller than $2s$, then we can be sure it is the unique most sparse solution to $Au = b$.

Unfortunately, determining the spark of a matrix is at least as hard as solving (P_0), so the stated condition is more of theoretical interest than practically applicable. A quantity which is simpler to compute is the *coherence* $\mu(A)$ of a matrix, defined as

$$\mu(A) = \max_{1 \leq j \leq N, i < j} \frac{\left| a_i^T a_j \right|}{\left\| a_i \right\|_2 \left\| a_j \right\|_2}, \tag{4.18}$$

where a_j is the jth column of A. $\mu(A)$ is also called the *mutual coherence*, but we will save this name for a slightly different measure introduced in subsection 4.3.2. The coherence leads to a lower bound on the spark: For a general matrix A we have

$$\text{spark}(A) \geq 1 + \frac{1}{\mu(A)}. \tag{4.19}$$

Combining this with the uniqueness condition involving the spark and further requiring that A have full column rank it is possible to derive the following result: Given a full column rank matrix $A \in \mathbb{R}^{M \times N}$ with $M < N$. If $Au = b$ has an s-sparse solution with

$$s < \frac{1}{2} \left(1 + \frac{1}{\mu(A)} \right), \tag{4.20}$$

then it is the unique solution to (P_1) as well as (P_0). The exact solution is also guaranteed to be found by OMP run with a termination threshold of 0.

This result is fascinating as it gives a condition for equivalence between (P_0) and (P_1): Essentially, under the stated conditions, we are guaranteed to find the most sparse solution (normally a task of combinatorial complexity) by solving (P_1) (solvable by efficient convex optimization algorithms). On top of that, under precisely the same conditions we have a form of equivalence between the convex relaxation approach of solving (P_1) instead of (P_0) and the very different greedy approach.

However, there is a price to pay for replacing the spark by the easier-to-compute coherence: While at first this result seems very promising, it turns out to provide very pessimistic guarantees on the sparsity of signals that can be recovered. We will give an example of that in the setting of CT in section 4.4.

The stated results involving the spark and the coherence are interesting because they address a general class of matrices. For any matrix, we can evaluate the sufficient condition for uniqueness using the coherence. However, we do not obtain any direct connection between the sparsity of a signal and the number of measurements we must acquire for recovering the original. Such a connection is the next topic, but it requires us to consider a restrictive class of matrices and accept results that rely on randomness.

We note that there are other measures of interest for characterizing equivalence of (P_0) and (P_1) and uniqueness of the most sparse solution, for example the exact recovery condition (ERC), [118], and the null space property (NSP), [50]. To the best of our knowledge, they also require a computational burden too large to be practically applicable [96].

4.3.2 Compressed sensing perspective

In a series of papers [22, 24, 25, 26, 27], Candès, Tao and Romberg together with Donoho [50, 49, 48] established what is now known as compressed sensing (CS), compressive sensing or compressed/compressive sampling. There are close connections with the sparse representation perspective but also some differences, most notably that randomness plays a key role, as we will show.

In CS it is assumed that the signal $u \in \mathbb{R}^N$ is represented in a orthonormal basis Ψ: $u = \Psi c$. Measurements are acquired as inner products with sensing vectors that are a subset of a full set that forms an orthonormal basis Φ. A measure of similarity called the *mutual coherence* is defined between vectors of the two bases:

$$\mu(\Phi, \Psi) = \sqrt{N} \cdot \max_{i,j} |\phi_i^T \psi_j|. \tag{4.21}$$

Note that the mutual coherence is computed using the full sensing matrix, not a particular sub-sensing matrix as in the sparse representation perspective. It can be shown that $\mu(\Phi, \Psi) \in [1, \sqrt{N}]$. If the mutual coherence is low, the two bases are called incoherent; if the lower bound is achieved they are called maximally incoherent, and if the upper bound is achieved, then they are called coherent. The interest in this kind of mutual coherence comes from results such as the following from [22], see also [28].

Given a signal u, assume that from the full set of measurements Φu only a *randomly* selected subset of size M is actually observed. Reconstruction through (P_1) recovers u exactly with overwhelming probability assuming that

$$M \geq C \cdot \mu^2(\Phi, \Psi) \cdot s \cdot \log N, \tag{4.22}$$

where C is a known and small constant. We see that the number M of required samples depends proportionally on the sparsity s, and only on the logarithm of the image size N, which means that potentially fewer measurements than N will suffice. It is also clear that the mutual coherence of the representation and sensing matrices is important, because only incoherent or low-coherent pairs admit an M smaller than N.

A natural question is what kind of measurement and representation matrices are incoherent? One example is the usual discrete Fourier transform and the pixel basis, which are in fact maximally incoherent, i.e., have a mutual coherence of 1. For this sensing-representation pair, the bound above reduces to $M \geq C \cdot s \cdot \log N$. This means that from any randomly selected subset consisting of (on the order of) $s \cdot \log N$ discrete Fourier samples it will be possible to recover, with high probability, the original image through solving (P_1). This is potentially a lot fewer than the full number of samples N needed in general to have a unique solution. On the other hand, for a coherent sensing-representation pair with mutual coherence of or close to $\sqrt{N}$ the bound says that more than N samples are needed, which is no improvement over naive reconstruction.

Another example of an incoherent sensing-representation pair is a sensing matrix with independent identically distributed (i.i.d.) elements, for example Gaussian, and some fixed representation matrix, for example simply a pixel basis. With high probability the mutual coherence will be close to $\sqrt{2 \log N}$. For example, for $N = 1000$, where the possible range for the mutual coherence is approximately $[1, 31.62]$, we have $\sqrt{2 \log N} \approx 3.71$, i.e., a low coherence.

Note that randomness has entered the picture, in fact in several places: The bound is concerned with measurements from a randomly selected subset of basis vectors from a sensing matrix. Alternatively, the sensing matrix itself contains elements randomly drawn from certain distributions. Also, the bound holds "with high probability", which means that it holds for the vast majority of signals but certain signals exist that violate the bound. Results that hold for *all* signals, not just with high probability, do exist, see e.g., [23], but then the $\log N$ is replaced by large powers, e.g., $(\log N)^6$, which gives a much larger bound on the required number of measurements than (4.22).

Another important concept in CS is the *restricted isometry property (RIP)*, which is defined in the following way: We say a matrix A satisfies the RIP of

order s if a constant $\delta_s \in (0,1)$ exists such that for all s-sparse signals u it holds that

$$(1-\delta_s)\,\|u\|_2^2 \leq \|Au\|_2^2 \leq (1+\delta_s)\,\|u\|_2^2\,. \tag{4.23}$$

As the name indicates, a matrix that satisfies the RIP of order s approximately preserves the Euclidean length of all s-sparse vectors, assuming the RIP-constant δ_s is not too close to one. The difference between two s-sparse signals can be up to $2s$-sparse, so if we require that A satisfy the RIP of order $2s$ then such signals can be distinguished from their measurements through A.

The RIP is useful for studying robustness, i.e., accurate reconstruction in the setting of noisy data, thanks to results such as the following from [25], see also [28]. There are two parts, the first assuming noise-free data:

If $\delta_{2s} < \sqrt{2}-1$, then the solution u_{P_1} to (P_1) from data $b = Au_{\text{orig}}$ satisfies

$$\|u_{P_1} - u_{\text{orig}}\|_2 \leq C_0 \cdot \|u_{\text{orig}} - u_s\|_1 / \sqrt{s}, \tag{4.24}$$

where C_0 is a constant and u_s is the original image with all elements, except the s largest, set to zero.

This result says that if the original signal is s-sparse, then it is recovered perfectly assuming the RIP-constant is small enough. Note that the claim is not probabilistic as (4.22) and does not explicitly require random sampling. Furthermore, if u is *not* s-sparse, the result puts a bound on the reconstruction error.

The second part of the result is for reconstruction from data subject to additive noise $b = Au + e$. Assuming the noise level is bounded $\|e\|_2 \leq \epsilon$ and that $\delta_{2s} < \sqrt{2}-1$, then the solution to problem (P_1^ϵ) in (4.16) satisfies

$$\left\|u_{P_1^\epsilon} - u_{\text{orig}}\right\|_2 \leq C_0 \cdot \|u_{\text{orig}} - u_s\|_1 / \sqrt{s} + C_1 \cdot \epsilon, \tag{4.25}$$

where C_1 is another constant. This is the same as the first result with the second term added, i.e., the reconstruction error is bounded by the sum of the error arising in the noise-free case and a constant multiple of the noise level. Again this result does not explicitly involve any randomness but holds for all signals as long as the RIP-constant is small enough.

However, randomness turns out to be relevant for the RIP-results after all, when trying to construct sensing matrices with low RIP-constants. Computing RIP-constants for a general matrix is NP-hard [96] so instead what has been shown are probabilistic results for certain sensing matrix classes with i.i.d. elements sampled for example from a Gaussian distribution: If the number of measurements satisfies

$$M \geq C \cdot s \cdot \log(N/s)\,, \tag{4.26}$$

then with high probability the matrix from such a class possesses the RIP, where C is a constant depending on which distribution the elements are drawn from. Also, such matrices with i.i.d. elements are in some sense near-optimal [27], since no other matrices can give the result (4.25) with fewer samples than (4.26).

4.4 Application to CT

There are different ways in which CT-images can be considered to be sparse: First, an image can be sparse directly in the image domain. This is for example the case in blood vessel imaging [81]. More commonly images are considered sparse in the gradient domain, i.e., after applying a discrete gradient operator across the image, since as described in section 3.3, CT-images tend to consist of larger relatively homogeneous regions separated by sharp boundaries. This matches well the use of TV as regularizer. An example of gradient image sparsity is given in Figure 4.5: The image is the classical Shepp-Logan phantom, see, e.g., [78], discretized onto a 256^2 pixel grid. It consists of 49% nonzeros, so it is not very sparse. The gradient image, on the other hand, is very sparse at only 3% nonzeros The use of the Shepp-Logan phantom is only to serve as an illustrative example; for a real image there will typically be a larger percentage of nonzeros. Other options for sparsity would be in a wavelet or frame such as curvelet or shearlet.

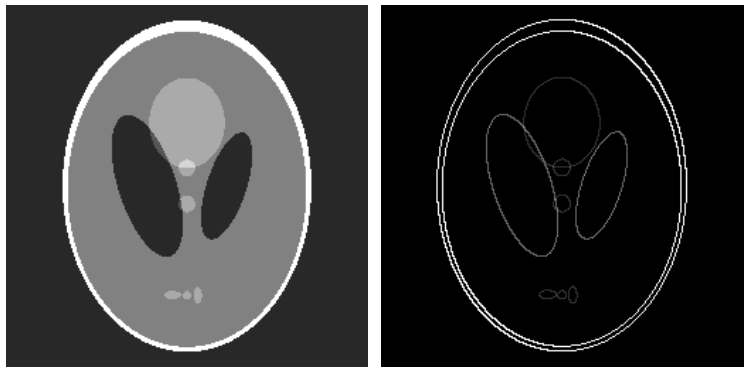


Figure 4.5: Left: The 2D Shepp-Logan test image. Right: The magnitude of the spatial gradient image. The image itself is not sparse but its gradient image is.

When attempting to apply the theoretical recovery guarantees from section 4.3 to the CT sensing matrix, namely the system matrix, as described in subsection 2.3.2, we quickly realize that neither the sparse representation or compressed sensing perspective are appropriate. From a sparse representation per-

spective, a CT matrix may not have full row rank, which prevents direct application of the coherence-test (4.20). Even if we ignore this aspect and compute coherence values (for CT matrices constructed using the software package AIR tools [71]), we find coherence values in the range 0.5 to 1 leading only to guaranteed recovery of 1-sparse images, see our paper C.

From the CS perspective there is no natural $N \times N$ "full-sampling" CT system matrix; the matrix can both have fewer and more rows than columns. The matrix is not orthogonal, nor does it have normalized rows. It does not contain randomly chosen rows and no i.i.d. elements, but deterministically computed elements based on the CT configuration. We are not aware of any results establishing useful RIP-constants for CT system matrices. A study of this was [107], in which the authors design a numerical strategy for determining a lower bound on the RIP constant of a CT matrix for a 512^2-pixel image. They find that already at sparsity $s = 100$, the lower bound is approaching 1. This means that only for images with sparsity substantially less than 100 there is hope for a small enough RIP constant to provide interesting bounds on the required number of measurements. This level of sparsity, in a 512^2-pixel image, is however much too low for most CT applications.

4.5 Summary

We have given an introduction to the field of sparse image reconstruction, including compressed sensing. We showed that sparse images can, under suitable conditions on the sensing matrix, be reconstructed accurately from fewer measurements than the number of unknowns, corresponding to an underdetermined linear system. Sparsity in the solution is promoted either by a greedy reconstruction method or through convex optimization, for example through 1-norm regularization. We presented a selection of theoretical bounds on the number of measurements needed to accurately reconstruct signals of a given sparsity, based on sensing matrix properties such as incoherence and the restricted isometry property. Finally, we addressed the application of sparse reconstruction to CT by giving the sparse reconstruction interpretation of TV-regularization and by showing that the theoretical recovery guarantees do not lead to useful bounds in CT.

Chapter 5

Characterization of sparse reconstruction in CT

In this chapter we discuss some remaining challenges for sparse image reconstruction in CT. In particular, we point to a fundamental lack of quantitative understanding of factors governing the achievable reconstruction quality. We argue that a set of computational characterization tools are needed for building such understanding. We present a general framework for analyzing sparse image reconstruction methods for CT to serve as the basis for developing such tools.

5.1 Practical challenges

As described in section 3.3, the potential for low-dose CT imaging through TV-regularization, and sparse reconstruction in general, have been demonstrated empirically. However, the mentioned empirical results have mainly taken the form of proof-of-concept and have not yet transitioned into routine use in for example medical CT. There are also examples of less successful application of TV-based image reconstruction to CT. Herman and Davidi report for example in [73] that TV-based reconstruction is unable to deliver useful reconstruction of their test image under realistic conditions. Similarly, in our own paper L investigating the optimal trade-off between few high-quality and many low-quality

projections for dedicated breast CT we found limited improvement by use of TV compared to an analytical FBP reconstruction.

Furthermore, similar to the denoising example in subsection 3.2.2, CT image reconstruction using TV-regularization have been criticized for being cartoon-like, "patchy", lose small details, as well as appearing unnatural in the eyes of radiologist that are used to the artifacts from analytical inversion methods, see e.g., [93, 114, 115].

It is not understood precisely what causes the different degrees of success of TV-regularization in the mentioned cases. As mentioned in section 3.3, there are somewhat vague suggestions that images must be "blocky" or piecewise constant for TV to work well, but no firm, quantitative conditions on which images are "sufficiently blocky" exist. CS is often used as motivation for CT image reconstruction from few projections, but as shown in section 4.4, CT does not match the assumptions of CS, so CS does not provide any quantitative explanation of the applicability of TV-regularized reconstruction in CT.

The lack of quantitative understanding leads to a large number of unanswered questions of high practical interest, for example:

- For what types of images (features, structure, contrast levels, etc.) can TV-regularization be expected to work well?
- What levels and types of noise, inconsistencies and artifacts can be suppressed by TV-regularization?
- How should the scanning configuration be designed to be best suited for TV-regularization?
- How to know whether the reconstruction quality obtained on a single test image will also be seen on other similar images?
- What other sparse reconstruction formulations are capable of yielding good reconstruction from few projections?

These questions and similar ones are, however, quite general, but they reflect a fundamental lack of understanding of factors that affect reconstruction quality, especially in a quantitative sense. We find that by adding to the general understanding it will be easier to provide quantitative answers to such questions when posed in specific contexts. In order to contribute such quantitative insight we find that a set of standard "characterization tools" must be developed.

The contributions of the thesis are aimed at developing an initial set of such characterization tools. The tools are to a large extent computational and empirically based, as theoretical and quantitative results for sparse reconstruction in CT are very limited. Before describing the specific contributions, we find that it will be useful to build a general framework of the various steps involved

in sparse reconstruction in CT. As we will see, such a framework allows us to analyze the posed questions more in detail and replace them by more specific questions that can be investigated systematically.

5.2 A general framework for sparse image reconstruction in CT

We see two general strategies for improving the fundamental understanding of sparse reconstruction in CT: The first strategy is to conduct extensive real-data studies for a specific application to get a thorough practical understanding of which factors affects the reconstruction quality. This strategy is very popular in the applied CT community, because it addresses the real-world case, which—in the end—is system that we want to improve. A limitation of this approach is that the system is treated as a black box in which it is difficult to isolate and study the role of individual parameters due to practical limitations of the system.

The second strategy studies sub-steps of the system to addresses basic questions in more idealized set-ups for establishing theoretical or empirical relations between for example noise level or image structures and reconstruction quality. This second strategy allows the researcher to abstract away from the complexity of data imperfections and focus on underlying relations. To be practically useful, however, conclusions made from idealized set-ups must be verified in real-data test cases. Our focus in the thesis has been on the second strategy.

Following this second strategy, we first construct a general framework of the interacting steps, domains and parameters involved in doing sparse image reconstruction for CT. By establishing such a framework, we will be able to more precisely understand the role of the different parameters. We will also use the general framework to specifically explain the thesis contributions.

The framework consists of four domains: The physical, the imaging model, the reconstruction model and the algorithm domains. Each domain has associated with it an image and a number of parameters as well as an error. An overview of the framework is shown in Table 5.1.

In CT image reconstruction we are interested in obtaining an image of an object which resides in the physical domain. We refer to this object as the *true image*. In the data acquisition step, it must be decided what data to record from the object, including what scanning configuration (parallel-beam/fan-beam/cone-beam etc.) to use, the x-ray intensity to use, and the number and positions

Domain	Image	Parameters	Image errors
Physical	True	Scanning configuration X-ray intensity Number of views Number of detector elements	-
Imaging model	Desired	Image expansion set Data discretization method Noise model	Model errors - systematic - stochastic
Reconstruction model	Designed	Regularizer Regularization parameter Constraints Optimization formulation	Design error
Algorithm	Algorithm	Choice of algorithm Internal parameters Initial point Stopping criterion	Algorithm error

Table 5.1: The general framework of sparse image reconstruction in CT. Each of the four domains has an associated image, a number of parameters and an image error with respect to the domain above it.

of the projections as well as the number and position of individual detector elements in a projection view. Some of these parameters may be fixed by the scanner available, for example in medical CT, the detector is fixed, while others can be decided when acquiring the data, for example the number of projections. These parameters all reside in the physical domain. We consider the physical image to be free from errors; it is the ideal image that we want to reconstruct. The errors associated with data acquisition are described in relation to the imaging model in the next step.

The next domain is the imaging model domain, in which a mathematical model is chosen to best describe the data acquisition process that takes place in the physical domain. The relevant mathematical models are the CC, CD, and DD imaging models from section 2.2, and we restrict our attention here to the DD imaging model, which is the relevant choice for sparse reconstruction methods. Within the DD model, the choice of expansion functions (e.g. pixels) and their number and the choice of how to discretize the data domain can be seen as imaging model parameters. We consider also the choice of the noise model as part of the imaging model parameters. Associated with each choice of imaging model parameters is a model error that describes "how far" the imaging model is from the physical domain. The model error consists of contributions from each model parameter, e.g., the choice of expansion set gives a *representation error*, and similarly there is a data-domain discretization error. Both of these are *systematic* errors in the sense that the model can be systematically improved for example by using a finer discretization. Measurement errors occurring in data acquisition are considered as inconsistencies with respect to the imaging model, and the part of the measurement errors that can not be explained by the chosen

noise model is classified as noise model error. In contrast to the systematic errors, the error arising from noise is *stochastic* as we cannot systematically modify the model to remove the error, since it is governed by the stochastic photon emission pattern of the x-ray source. Given a set of data, the *desired image* is the image which gave rise to the data through the specified imaging model.

As described in subsection 3.1.2, the imaging model can be ill-posed as well as have an ill-conditioned system matrix A, which makes it difficult or impossible to determine the desired image. In the third domain, a reconstruction model is designed to handle these problems. The reconstruction model is an optimization problem which is set up to provide regularization of the imaging model. The possible choices of regularizers, regularization parameter, constraints and optimization problem formulation are the parameters of the reconstruction model domain, the *design parameters*, that determine the *designed image*, i.e., the mathematical solution of the optimization problem. In general, the choices of design parameters lead to a *design error*, i.e., a difference between the designed and the desired solutions. Ideally, the optimization problem should be chosen so as to have a the smallest possible design error.

The final domain is the algorithm domain, which is entered when in practice we must solve the reconstruction model numerically. To do that, we must use a numerical algorithm, and depending on the choice of algorithm we may or may not be able to achieve the designed solution. We call the output from the numerical algorithm the *algorithm image*, and the difference between the algorithm image and the designed image we refer to as the *algorithm error*. Associated with any algorithm is a number of internal *algorithm parameters* that specify for example step lengths, the initial point and termination criterion, that affect the algorithm error. The choice of algorithm itself can also be considered an algorithm parameter. However, the algorithm parameters *only* affect the algorithm error, not the design error, which is determined by the design parameters in the reconstruction model domain. An algorithm error can arise from the use of an optimization algorithm, which is not convergent to the minimizer of the optimization problem, or from terminating a convergent algorithm early, before convergence is reached.

We find that thinking about reconstruction in terms of this general framework makes it possible to better distinguish the different phases and goals and their connection. We can now dissect the questions from section 5.1 further into more specific questions. The following subsections analyze different aspects of reconstruction by means of the general framework. The general framework is also used in the next chapter when describing the contributions of the thesis.

5.3 Experimental design issues

5.3.1 Inverse crimes

When working with inverse problems, in particular in simulation studies, there is a caveat to be aware of. It is referred to as an "inverse crime", see, e.g. [77, 86]. As inverse crime occurs when precisely the same model is used for generating simulated data and for reconstructing an image from that data. For example in CT, if synthetic discrete data is generated as $b = Au$ from a discrete image u, then also using this same A in the reconstruction is an inverse crime. The problem is that no model errors are present in the simulation, which tends to lead to overly optimistic reconstruction quality compared to what can be obtained in a more realistic set-up, such as when reconstructing from real data.

We can interpret an inverse crime in terms of the general framework. Within the DD imaging model, synthetic data is computed, not from the true image in the physical domain, but from its expansion coefficients. In the reconstruction stage, the same set of coefficients are determined, and from ideal noisefree data it can be possible to recover the coefficients perfectly. This, however, does not address the question of the representation error from using for example a pixel expansion of the true object.

Some of the initial results of CS that spurred immense interest by promising *exact* reconstruction, see e.g., [24], were in fact, inverse crime studies, since the same DD imaging model was used for data generation and reconstruction. This does not mean that exact reconstruction is not surprising or worth studying. It is simply important to realize that the reconstruction is only perfect up to the discrete representation to avoid too high expectations when going to more realistic scenarios.

One way to avoid the inverse crime in simulation with synthetic data is to use different representations, for example different pixel grid sizes, for the synthetic data generation and reconstruction. While it is not the same as acquiring data from the physical true image, this approach introduces a model error in the reconstruction relative to synthetic data and thus prevents the inverse crime.

However, we find that there is something to learn from deliberately doing inverse crime studies. First of all, inverse crime results can be a good proof-of-concept when initially working with a new reconstruction problem. But just as important, we find that inverse crime studies are needed when attempting to characterize reconstruction problems. Model errors contribute to complicating the imaging problem and by focusing on the inverse crime scenario, the

important factors affecting for example the design error can more easily be characterized. With insights obtained from inverse crime characterization, the non-inverse crime scenario can be addressed by including modeling errors.

5.3.2 Assessing image quality

A fundamental question in reconstruction is: how closely does the reconstructed image (the algorithm image in the general framework) approximate the underlying true image in the physical domain? Ideally, we hope that the reconstructed image reproduces the true image perfectly, but in practice we cannot expect this to happen. Let us refer to the error of the algorithm image w.r.t. the true image by the *reconstruction error*. How can we measure the reconstruction quality so that we can, for example, compare different reconstructions and tell which one is better? The question is complicated enough if we assume knowledge of the true image for use as reference, but even more complicated in the realistic case of an unknown true image. We consider first the simpler case of the true image being known.

The general framework tells us that the reconstruction error consists of model, design and algorithm errors. With the DD imaging model, the true image is replaced by a discrete representation (the desired image) with an associated model error, more specifically a representation error from representing the image by the finite expansion set. Assuming knowledge of the desired image, the simplest choice for quantifying reconstruction error is to use general mathematical image quality metrics such as the root-mean-square error (RMSE), signal to noise ratio (SNR), and the structural similarity index (SSIM) [125]. These choices can be considered to measure a distance between the algorithm image and the desired image, so that the best reconstruction occurs at a measure of zero. The advantage is conceptual and computational simplicity. The drawback, however, is that in general there is no direct connection between these distances and the quality of an image. As argued by Barrett and Myers [4], image quality assessment must take into account the specific *task* that the reconstruction is designed to solve. If this is not done, then one method may appear better than another method in terms of the metric, while in fact the improvement as measured by the metric was caused by improvements irrelevant for the task. Furthermore, ignoring the representation error only measures the reconstruction error relative to the (discrete) desired image. This corresponds to an "inverse crime", as discussed in subsection 5.3.1.

A more appropriate method for relevant image quality assessment is observer studies. If for example, the task of a specific reconstructed image is to be interpreted by a radiologist for detecting tumors, an observer study can be

designed in which reconstructed images from different methods are evaluated by a panel of radiologists for that specific task. The drawbacks, however are clear: observer studies are time-consuming and reproducibility can be difficult due to biases held by the observers. A compromise between mathematical metrics and observer studies is to use observer models [5], in which a mathematical model of observers is built to capture some characteristics of human observers while allowing for evaluation through computation.

We stress that specific choices of image quality metrics must be based on what goal we are trying to achieve. In most of our studies we are interested in the algorithm error or the design error in order to characterize reconstruction methods under ideal conditions. For that purpose we find that the use of simple metrics such as the RMSE is appropriate. However, when generalizing to more realistic scenarios, and the practical utility of the reconstructed image becomes the goal, then attention must be payed to the challenge of image quality assessment.

5.3.3 Comparing reconstruction models

When studying different reconstruction models to determine the better one for a particular imaging application, the general framework tells us that the relevant quantity to study is the design error. The design error only depends on the reconstruction model used and should therefore not be affected by the choice of optimization algorithm. In practice, to enable a fair comparison of reconstruction models, this calls for running a convergent optimization algorithm for sufficiently many iterations that the algorithm error is very close to zero for all reconstruction models in question.

How fast the optimization algorithm converges is in principle not relevant either, as long as eventually the designed solution is arrived at. At a later stage, when a particular reconstruction model has been selected, one can start looking for the fastest algorithm. For running such comparative studies between reconstruction models, it is useful to have a general algorithm that can be adapted to new optimization problems with little implementation effort. We consider such a "prototyping" or "Swiss army knife" optimization algorithm in our paper E.

5.3.4 Comparing algorithms

When choosing which numerical algorithm to use in a given specific application, there are several aspects to pay attention to. First of all, one possibility is to use a heuristic algorithm not proven to be convergent for obtaining an algorithm

image which—in some sense—approximates the designed image. This, however, undermines the distinction of the general framework between the desired, designed and algorithm solutions, since if the algorithm is not convergent to the designed image, it makes more sense to ask what is the difference between the algorithm and desired solutions directly.

By demanding a convergent algorithm, we can split that question in two and study *separately* how closely the designed solution approximates the desired and how close the algorithm solution is to the designed. Further, it makes sense to compare convergent algorithms in terms of how fast they converge to the designed image.

The choice of termination or convergence criterion is important here. An often seen strategy in inverse-crime studies is to monitor the error of the algorithm image with respect to the original (desired) image. Since in general the designed solution is different from the desired, this strategy will not observe the algorithm-to-desired image error approach zero but instead the constant designed-to-desired image error. If the goal is to compare algorithms in terms of speed, we find it more meaningful to monitor the algorithm-to-designed image error, i.e., the algorithm error of Table 5.1.

5.4 Summary

In this chapter, we identified a need for improving the quantitative understanding of conditions under which sparse reconstruction can be expected to perform well in CT. We proposed to develop a set of characterization tools for establishing such fundamental insight. To that end we introduced a general framework for systematically analyzing and describing sparse reconstruction methods and used it to discuss comparison of optimization algorithms, aspects of image quality assessment and the role of inverse crime simulation studies.

Chapter 6

Contributions

This thesis is aimed at adding to the fundamental understanding of the application of sparse image reconstruction methods to CT. The contributions can be seen as an attempt towards bridging the gap between the CT community and the mathematical imaging community. The broad theme in the contributions is development of computational tools (algorithms, software, techniques) for characterizing sparse image reconstruction for CT.

Our studies can be broadly divided into two major groups. First, a number of studies are focused on development of optimization algorithms for TV-regularization and other sparse reconstruction problems. While some emphasis is put on fast, efficient algorithms, the main focus here is on "prototyping" algorithms, i.e., general-purpose algorithms and software for *reliably* solving various sparse reconstruction models. This work is of interest in its own right and it provides an algorithmic foundation for the central aim of the thesis: *Characterization* of sparse image reconstruction for CT, which the second set of studies are focused on. By characterization, we mean establish quantitative relations between parameters of the general framework in Table 5.1. More specifically, we focus on the role of image sparsity and study connections between sparsity and the required number of projection views for accurate reconstruction.

Here, we give here a chronological overview of the thesis contributions to illustrate the development of our motivation to pursue the two major research

directions of prototyping algorithms and characterization methods. The relevant papers are described in turn with the most significant results highlighted. The contributions are put into context of the general framework in Table 5.1. For more detailed description of the contributions we refer the reader to the papers enclosed in the appendices.

6.1 Initial motivation: TV-regularization for tomographic reconstruction using accelerated first-order methods

In the initial phase of the PhD project, the focus was on an application of tomographic reconstruction in materials science. Specifically, we were interested in improving the reconstruction quality in imaging of polycrystalline materials, such as metals, from x-ray diffraction data. The data corresponded to 2D parallel projections of individual 3D crystals (called grains) in a sample. The reference reconstruction method used was ART [84, 98]. The sharp edges between the polyhedral crystals and the background and the fact that relatively few projections were recorded motivated us to apply TV-regularization with the hope to improve the ART reconstructions.

Motivation and goals:

Our first goals were to develop simulation software for the tomographic reconstruction problem and efficient optimization algorithms and software for solving the TV-regularized reconstruction problem, as no suited existing software was found. In particular, while software for TV-based 2D image denoising and deblurring were available in the public domain at the time (2009), our large-scale 3D application called for new and highly efficient algorithms.

Relevant papers:

[F] Implementation of an optimal first-order method for strongly convex total variation regularization (page 213).
[K] Accelerated gradient methods for total-variation-based CT image reconstruction (page 271).

Methods and results:

For the tomographic simulation software part, we designed a set of MATLAB® functions to set up and run numerical experiments with the system matrix

of the particular tomographic reconstruction problem. This work resulted in a software package, called `tomobox`, which is available from `www.mathworks.com/matlabcentral/fileexchange/28496-tomobox`. Further, the work also led to the development of MATLAB® functions implementing parallel-beam and fan-beam CT geometries, which were published in [71] as part of the software package AIR Tools, available from `www.imm.dtu.dk/~pcha/AIRtools`.

We focused our optimization algorithm work on implementing a selection of first-order optimization algorithms for the TV-regularized least-squares problem with Huber-smoothing of the TV-term (3.26). The most important, and novel, algorithm implementation was a particular accelerated first-order algorithm proposed in the theoretical setting by Nesterov [90, 91]. The algorithm assumes knowledge of the objective function's Lipschitz constant and strong convexity parameter, which in general are unknown. In the implementation we used several heuristics for locally estimating these parameters. Our theoretical analysis of the heuristics-equipped algorithm showed that it has a near-optimal convergence rate, compared to the theoretical algorithm. We named our algorithm UPN (Unknown Parameter Nesterov).

We compared the performance of UPN with three other algorithms: 1) the standard gradient method, 2) a variant [103] using the Barzilai-Borwein (BB) step-size selection [6] with a non-monotone line search [65], and 3) a version of the accelerated Nesterov method that does not use the heuristic for the strong convexity and is optimal for the class of non-strongly convex problems. The BB algorithm has been shown empirically to be much faster than the standard gradient method but it does not belong to the optimal class. The algorithm implementations in MATLAB® were collected into a software package, called `TVReg`, and made available from `www.imm.dtu.dk/~pcha/TVReg`. The numerical tests were performed on synthetic 3D tomographic reconstruction test problems generated by the `tomobox` software.

We found that for large values of the Huber smoothing parameter the methods were comparable. This case can be thought of as "easy" due to heavy smoothing, and as expected low iteration numbers were observed for convergence. For smaller values of the Huber smoothing parameter and even more so with larger TV regularization parameters (the "harder case"), UPN outperformed the other methods, particularly in the case of requiring a high-accuracy solution.

We also compared observed convergence rates on CT image reconstruction problems with system matrices corresponding to an overdetermined and an underdetermined system of linear equations. Here, we observed a comparable performance of the BB-algorithm and UPN on the overdetermined case, while in the underdetermined case UPN severely outperformed the BB-method. This was interesting for two reasons: First, the underdetermined case corresponds to a

low-dose CT imaging scenario and the fast convergence of UPN is promising for obtaining a practical algorithm for this important case. Second, in the underdetermined case, the optimization problem does not satisfy the assumption of strong convexity underlying UPN. The observed fast convergence indicates that the local heuristic employed by UPN for estimating the (global) strong convexity parameter can deliver the desirable faster convergence rate for a *non*-strongly convex problem.

In the general framework of Table 5.1, this work belongs in the algorithm domain, as the focus is on the algorithmic aspects such as rate of convergence, rather than on the image quality of TV-reconstructions. The algorithms developed in this work are applicable for solving TV-regularized problems in CT of fairly large size and have been used extensively as "work horse" for a number of other projects, including some in this thesis.

Conclusions:

- We achieved the goal of constructing software for both tomographic reconstruction and efficiently solving the TV-regularized reconstruction model.
- For observed convergence rates, we found large dependence on both the TV regularization parameter and the Huber smoothing parameter (both design parameters, *not* algorithm parameters), causing different algorithms to be faster on different test cases.
- We also saw a large discrepancy between the observed convergence rates and the theoretical worst-case rates, teaching us the lesson that worst-case convergence rates do not necessarily tell which algorithm is faster in practice.

With the software in place, the natural next step was to apply it to the materials science tomography problem to investigate whether TV-regularization would lead to improvements over ART. For a number of different reasons, this work was never realized. Most notably, the required x-ray diffraction data never became available. Instead, the focus of the project was shifted into application of TV-regularization in medical CT.

6.2 Gaining experience: Empirical studies of TV-regularized tomographic reconstruction

The application of interest in the next phase of the project was low-dose CT scanning for breast cancer screening, as explained in subsection 3.3.1. Again, our goal was to demonstrate reconstruction improvements through the use of

TV-regularization; this time over the filtered back-projection (FBP) analytical reconstruction method.

6.2.1 The optimal distribution of a fixed x-ray exposure?

Motivation and goals:

In this study, we addressed the question of how to best distribute a fixed total x-ray exposure among the projection views. A higher x-ray exposure gives a better signal-to-noise ratio, but in the application of screening for breast cancer, it is imperative to keep the exposure low to reduce the risk of radiation-induced cancer. Therefore, given the maximal allowable exposure (comparable to the expose used in mammography), the question was how to distribute the exposure to get the best possible reconstruction. Essentially, the choice is between acquiring many high-noise projections (of low quality) or fewer low-noise projections (of higher quality), and our goal was to determine the optimal trade-off.

Relevant papers:

[L] Toward optimal X-ray flux utilization in breast CT (page 277).

Methods and results:

We designed a simulation study to investigate the trade-off using a simulated 2D breast phantom image as a case study. The breast phantom had semi-realistic tissue distribution and of particular focus was to reconstruct of a group of tiny high-intensity calcium deposits, known as microcalcifications, which are a known early indicator of a developing cancer. We did two sets of experiments: One with noise-free data and one with simulated Poisson noise, keeping the total x-ray intensity fixed while using different numbers of projection views. In this way we obtained different cases ranging from very few projection views of good quality to many projection views of low quality. We compared reconstruction quality by TV-regularization and FBP.

In case of noise-free simulations, we found, as expected, that more projections lead to better reconstructions and that TV-reconstruction can produce slightly more well-resolved microcalcifications than FBP. In case of noisy data, we found that both TV and FBP reconstructions had difficulty in faithfully reconstructing the microcalcifications. The TV reconstructions suffered from the well-known patching/staircasing artifacts and/or a cartoonish appearance, while the FBP reconstructions were contaminated by large amounts of noise. From visual inspection it appears that many high-noise views give slightly better reconstruc-

tion. This is interesting, because many studies show (as described in subsection 3.3.2) that accurate TV-reconstruction is possible from a reduced number of projection views, but without a meaningful image quality assessment method we cannot make any firm conclusion.

In the general framework of Table 5.1, this study has elements in the physical and the reconstruction model domains, because parameters of the physical scanning configuration are being studied at the same time as the design error.

Conclusions:

- We were not able to obtain a clear conclusion on the best trade-off between few good-quality and many low-quality projections.
- We did not succeed in demonstrating improvements by TV-regularized reconstruction over FBP.
- Maybe TV-regularization is not the best choice of reconstruction model for this application. Could other choices give better reconstruction?

6.2.2 Non-uniform convergence issues

Motivation and goals:

In this work, we focused on a particular algorithm-oriented challenge we encountered while working with TV-regularized reconstruction in the breast cancer screening application. In general, ideally, we want the algorithm to return the designed solution to the TV reconstruction model, but this may require an impractically long computational time. In practice, we want to run the algorithm long enough that the algorithm solution is "sufficiently close" to the designed solution, but not longer to avoid unnecessary waiting time. Determining the optimal termination point was the motivation for this study.

Relevant papers:

[J] Ensuring convergence in total-variation-based reconstruction for accurate microcalcification imaging in breast X-ray CT (page 265).

Methods and results:

In a simulation with TV-regularized reconstruction of the breast phantom image with microcalcifications, we observed the algorithm solution iterates approach the designed solution in a very non-uniform way: the background breast tissue was quickly approximated closely, but the pixel values within the high-

intensity microcalcifications were subject to much slower convergence. Further we showed that a standard global convergence criterion, such as the norm of the objective function's gradient being close to zero, has difficulty in indicating the non-converged values. Instead we proposed to monitor the complete gradient of the objective function as an image. We showed that this strategy allows better detection of non-converged pixel values and demonstrated improvement in ensuring a globally converged image.

In the paper, only convergence to designed image was studied. The important question of design error was not touched. We observe that even when fully converged to the designed solution, the intensity of the reconstructed microcalcifications is much lower than the intensity in the original image. This intensity reduction is a well-known model error, or bias, of TV-regularization, see subsection 3.2.2. In the case of microcalcifications, this effect is particularly pronounced due to their small scale, see [113].

In the general framework of Table 5.1, this work belongs in the algorithm domain, as the focus is on the algorithm-specific question of ensuring convergence. However, the problem of the intensity reduction model error is an aspect of the reconstruction model domain.

Conclusions:

- We illustrated highly non-uniform convergence of pixel values to the designed TV solution, which cause the sufficiently accurate reconstruction of microcalcifications to be a challenge.
- We proposed to monitor gradient components of the objective function to better detect non-uniform convergence, but did not construct a specific termination criterion.
- We noted a large model error of the designed solution manifested as intensity reduction of the microcalcifications. This lead us to ask, once again, whether TV-regularization is a good choice for the particular application, or other choices might give a better result.

Our general impression from empirical studies with TV-regularized for tomographic reconstruction was—in a single word—disappointment. Based on the literature, we had expected large improvements over established reconstruction methods such as ART and FBP, but were only able to demonstrate very minor improvements, if any at all. Moreover, it was not clear *why* this was the case. *What* caused TV-regularized to perform poorly in our case studies? For example, the breast image contained sharp edges and was very sparse in the gradient domain, commonly stated reasons to expect TV-regularization to work well. But on the other hand, the breast tissue structure was somewhat complex, not "blocky", as also mentioned as criterion for good TV-reconstruction [123].

Dissatisfied by such qualitative explanations, we were motivated to investigate *quantitatively* which factors affect the achievable reconstruction quality. This motivated a line of work on *characterizing* TV-reconstruction, and more generally, sparse reconstruction for CT, i.e., establish quantitative relations between the various parameters of the general framework in Table 5.1. In particular, we were interested in quantifying the role of image sparsity, as motivated by the theoretical guarantees from compressed sensing.

At the same time, we asked ourselves whether other reconstruction models than TV could work better. In order to experiment with other reconstruction models, we needed to have optimization algorithms for reliably determining the designed solution to avoid seeing effects of inaccurate algorithm solutions. We wanted to avoid the extensive work required to develop dedicated convergent optimization algorithms for each reconstruction model of interest, as it was done for TV-regularization in the `TVReg` software of paper F. This need motivated us to initiate a line of work on *prototyping* optimization algorithms, i.e., an optimization framework with a convergent algorithm generally applicable to a wide range of reconstruction models of interest in CT. Our work on prototyping algorithms is described next, before turning to our work on characterization.

6.3 Prototyping algorithms: Tools for comparing reconstruction models

6.3.1 The basic prototyping algorithm: Chambolle-Pock

Motivation and goals:

In this work, we were motivated to obtain an optimization algorithm generally applicable to a variety of reconstruction models and fast enough to produce designed solutions in reasonable time. In other words, we wanted to develop an optimization algorithm framework for prototyping reconstruction models for CT image reconstruction. With such a framework, constructing a convergent algorithm for a new reconstruction model of interest becomes a simple, mechanical task.

Relevant papers:

[E] Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm (page 183).

Methods and results:

We worked with a primal-dual optimization algorithm introduced by Chambolle and Pock (CP) [31]. As the algorithm is applicable to a very broad class of general imaging problems, we found it well-suited to serve as basis for a prototyping algorithm. Depending on the smoothness of the problem, different variants of the CP algorithm can be applied, with favorable worst-case iteration complexity for smooth problems. An important aspect of the algorithm is that it can handle non-smooth problems, which allowed us to construct algorithms that avoid smoothing the TV.

From the generic CP algorithm, we derived specific algorithm instances for reconstruction models of interest in CT image reconstruction, for example the TV-regularized least-squares problem, the data-constrained TV-minimization problem and less-studied problems such as involving the Kullback-Leibler (KL) divergence and the 1-norm for the data fidelity term. Deriving each algorithm instance was to a large extent a mechanical step-by-step procedure. We verified the derived CP algorithm instances by running simulations of a CT image reconstruction test problem until convergence to the designed solution was seen. All derived algorithm instances were found to be convergent. We therefore considered the CP algorithm well-suited for reconstruction model prototyping.

In addition to the specific example algorithm instances we gave, the paper is intended as a "cookbook" for deriving CP algorithm instances for other optimization problems of interest to a CT engineer.

In the general framework of Table 5.1, the work done in the paper is within the algorithm domain, as it focuses on deriving specific algorithm instances. By using the CP algorithm for prototyping optimization problems, it becomes possible to do studies in the reconstruction model domain, i.e., compare the designed solutions of various reconstruction models for a given imaging application.

Conclusions:

- We built and demonstrated an algorithmic framework for prototyping reconstruction models using the Chambolle-Pock primal-dual optimization algorithm.
- We constructed a number of algorithm instances for reconstruction models relevant for CT and demonstrated convergence.
- Put in a different way, we have communicated the Chambolle-Pock algorithm from the mathematical imaging community (in a rather technical article [31]) to the CT community as a practical, flexible and easily applicable optimization algorithm and tool for prototyping reconstruction models.

After having constructed the algorithmic prototyping framework, we wanted to illustrate its flexibility by demonstrating the use in multiple non-standard reconstruction models. The first such work extends the prototyping framework to reconstruction models based on convex feasibility problems.

6.3.2 Convex feasibility paper

Motivation and goals:

We proposed the use of convex feasibility (CF) problems as reconstruction models for CT image reconstruction as an alternative to optimization problems with a unique solution. A CF problem consists of a number of convex constraints that a solution must satisfy. The solution set is the whole feasible set, i.e., the intersection of the constraint sets. There is no objective function to select a specific solution within the feasible set. Therefore there can be multiple equally good solutions, or in the case of an empty intersection, there is no solution.

Relevant papers:

[B] First-order convex feasibility algorithms for x-ray CT (page 105).

Methods and results:

We adapted the CP algorithm from paper E for CF problems, or more precisely for a modified CF problem, where the constraints were incorporated as indicator functions into an objective function. The objective function contained an additional squared 2-norm penalty measuring deviation of the solution from a prior image in order to select specific solutions among the feasible solutions. We derived specific CP algorithm instances for CF problems with data equality constraints and inequality constraints, including an instance for a CF problem involving TV-regularization.

In numerical simulations we verified that the derived algorithm instances were indeed convergent on a limited-angular range CT test problem. We compared the basic CP algorithm for non-smooth problems with an accelerated version for smooth problems and as expected the accelerated version was considerably faster. The existence of a fast method is particularly important for limited angular range CT problems that are known to be more difficult to solve accurately due to larger condition number of the system matrix than for full-angular range CT.

In short, the paper served to exemplify the use of the CP algorithm for CF problems in the same "cookbook"-style as paper E. As such, this work also belongs in the algorithm domain of the general framework of Table 5.1. By using derived CF algorithm instances it will be possible (in a future work) to conduct the reconstruction model domain study of prototyping various CF reconstruction models for use in a particular application.

Conclusions:

- We extended the prototyping framework to convex feasibility problems as an alternative to reconstruction models with a unique solution.
- We derived and verified the convergence of several algorithm instances for convex feasibility reconstruction models.

While the convex feasibility extension of the prototyping framework was another algorithm domain work, we addressed in our next work the intended goal of prototyping reconstruction models for a particular reconstruction.

6.3.3 Reconstruction model prototyping for SPECT

Motivation and goals:

Our goal with this work was to study a particular reconstruction model, based on a blurred piecewise constant object model, for its potential to improve reconstruction quality from few-view data in single photon emission computed tomography (SPECT). SPECT inherently produces blurred low-resolution images, which motivated the use of a blurring operator in the reconstruction model.

Relevant papers:

[D] Few-view single photon emission computed tomography (SPECT) reconstruction based on a blurred piecewise constant object model (page 143).

Methods and results:

We derived a specific CP algorithm instance for a reconstruction model involving TV-regularization of a deblurred version of the image. Using the CP algorithm instance, we conducted an extensive set of simulations to assess the potential of the proposed reconstruction model. Compared to the relevant reference reconstruction method, the maximum-likelihood expectation-maximization (MLEM) algorithm, we found that the proposed reconstruction model shows convincing potential for improved reconstruction from few projection views.

We emphasize that the important aspect of this work—in relation to the present thesis—was the straightforward construction of the CP algorithm instance that made it easy to obtain the designed solution of the proposed reconstruction model. That enabled our expert SPECT collaborators to thoroughly assess the proposed reconstruction model, without having to do any optimization algorithm development work themselves.

Since this work studies properties of the designed solution, it belongs in the reconstruction model domain of the general framework of Table 5.1; the prototyping algorithm framework is only used as a means of getting to the designed solution.

Conclusions:

- We derived a prototyping algorithm instance for the proposed reconstruction model for SPECT.
- Using the algorithm instance, we demonstrated the potential for improved reconstruction quality from few-view data with the proposed reconstruction model.

Our final prototyping work was another case study of a specific reconstruction model's potential for improving reconstruction quality.

6.3.4 A non-convex reconstruction model

Motivation and goals:

We wished to investigate the possibility of improved CT reconstruction by the use of a nonconvex variation of TV-regularization. Specifically, we replaced the 1-norm of the TV-regularizer by a "p-norm" with $p < 1$ as in (3.25). The goal was to demonstrate that the nonconvex problem leads to greater sparsity-promotion and therefore the capability to reconstruct images sparse in gradient more accurately, i.e., with smaller design error, than the standard convex TV.

Relevant papers:

[H] Nonconvex optimization for improved exploitation of gradient sparsity in CT image reconstruction (page 249).

Methods and results:

We derived a prototyping algorithm instance for the non-convex TV-regularized

problem based on the convex feasibility formulation. In order to handle the non-convexity we equipped the algorithm with a set of iteratively updated weights.

We carried out a simulation study to assess the reconstruction quality obtainable by the proposed method. We used two different test images, one of them is the breast phantom image with microcalcifications for which we earlier observed unsatisfactory convex-TV-regularized reconstructions. Using $p = 0.8$ we saw significantly improved reconstruction at few projections over standard $p = 1$ TV.

As the goal of this work is to study quality of the designed solution to the nonconvex reconstruction model, it belongs in the reconstruction model domain of the general framework in Table 5.1.

Conclusions:

- We derived a prototyping algorithm instance for a non-convex TV reconstruction model.
- We demonstrated the potential for improved reconstruction over convex TV-reconstruction.

To summarize our prototyping work, we can say that through the first two "cookbook" papers and the following two case studies, we demonstrated the usefulness of the algorithm framework for reconstruction model prototyping. We are in a position to easily carry out further prototyping studies of reconstruction models, either in our familiar frame of CT, or together with collaborators working on other applications of tomographic imaging.

Furthermore, with the prototyping framework in position, we are now equipped with an algorithmic tool to study relations between parameters of sparse reconstruction methods by systematically executing a large number of simulations with various sparsity-based reconstruction models. This is our work on characterization that we describe next.

6.4 Characterization: Tools for analyzing sparse reconstruction in CT

In our characterization studies we were interested in improving the general understanding of conditions under which sparse reconstruction methods lead to good reconstruction quality. More specifically, we wanted to establish quantitative relations between parameters of the general framework in Table 5.1.

Motivated by CS results connecting image sparsity and sufficient sampling for accurate reconstruction, we focused on investigating the role of sparsity in CT. Due to the lack of theoretical guarantees extending to CT, we took an empirical approach.

6.4.1 Quantifying undersampling

Motivation and goals:

Motivated by CS results proving a relation between image sparsity and the number of measurements sufficient for accurate reconstruction, we wanted to investigate the existence of a similar relation in CT. As described in chapter 4, CS establishes quantitative results on what levels of undersampling can be tolerated in reconstruction based on 1-norm minimization, but for certain matrices only, and not including CT system matrices. One particular difference of CT system matrices is the lack of a natural full-sampling point, corresponding to $N \times N$ matrices used in standard CS. A square CT system is not necessarily of full rank, nor well-conditioned, and hence does not specify a unique solution. We addressed the question of establishing an equivalent full-sampling point to serve as reference for quantifying the undersampling that can be attained by sparse reconstruction methods.

Relevant papers:

[A] Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in x-ray CT (page 89).
[I] Sampling conditions for gradient-magnitude sparsity based image reconstruction algorithms (page 255).

Methods and results:

We proposed and discussed four different *sufficient sampling conditions* (SSCs) as candidates for such a reference point. All conditions can be expressed in terms of the singular value decomposition of the system matrix. We introduced the first SSC, referred to as SSC1, to simply correspond to a square matrix, and demonstrated that SSC1 does not guarantee invertibity. The next, SSC2, corresponds to taking precisely enough projection views and number of detector elements that the system matrix has full rank. SSC2 guarantees a unique solution but does not address Hadamard's stability condition (page 20). Stability can be studied in terms of the system matrix condition number, and when increasing the number of projections and detector bins, the condition number decays toward a limiting value. We defined SSC3 to describe the number of

projections and detector bins that cause the system matrix condition number to be within a certain neighborhood of the limiting value. SSC3 requires computation of the smallest singular value of A, which can be very time-consuming. For the considered full-angular range fan-beam CT configuration, we introduced SSC4 as $2N_s$ projections and $2N_s$ detector bins for a disk-shaped discrete image inscribed in an $N_s \times N_s$ array as a simple heuristic for SSC3. While the specific CT configuration considered is the full-angular range fan-beam configuration, SSC2 and SSC3 apply to other configurations as well.

We studied properties of the proposed SSCs and showed that similar discretization methods for the DD imaging model lead to similar, but slightly different, SSCs. We also showed that the SSCs scale linearly with image size.

We used the proposed SSCs as reference points of full-sampling for studying the levels of undersampling that can be achieved by reconstruction through data-constrained TV-minimization (3.23). As a case study we used the breast phantom with microcalcifications also used in paper L and paper J. We studied both undersampling in number of projections, for a fixed number of detector bins per projection, and undersampling in detector bins, for a fixed number of projections. In contrast to data-constrained variants of Tikhonov regularization for which on the order of SSC2 projections were needed for accurate reconstruction, we found the TV-reconstruction error to decay rapidly to an almost constant low level, establishing an admissible projection undersampling of about a factor of 2 relative to SSC2 for accurate reconstruction. Interestingly, for detector bin undersampling, we observed a more gradual TV reconstruction-error decay and concluded that detector bin undersampling was not possible. Similar observations were made with respect to stability.

We did a preliminary study of gradient sparsity influence on the admissible undersampling by comparing results for test images with different sparsities. For a reduction in gradient sparsity of a factor 4, i.e., a test image of much less complex structure, we saw a gain of the same factor 4 in undersampling admitted by TV-reconstruction. This indicates a simple relation between sparsity and sufficient sampling for accurate sparsity-exploiting image reconstruction.

We repeated the study using qualitatively different test images with the same sparsity as the considered breast phantom image. We found similar undersampling levels attainable through TV-minimization, which further supported our conjecture of existence of a sparsity-sampling relation.

In the general framework of Table 5.1, this work belongs partly in the imaging model domain and partly in the reconstruction model domain. The proposed SSCs are properties of the imaging model in question and our studies address how the SSCs change with variations to the DD imaging model. When using the

SSCs as reference point for full-sampling for quantifying undersampling by TV-reconstruction, we are in the reconstruction model domain because our interest is in properties of the designed solution.

Conclusions:

- We proposed four sufficient sampling conditions (SSCs) to serve as reference point for quantifying admissible undersampling by for example TV-based image reconstruction.
- We showed that the proposed SSCs scale in a simple way with the image size, which means that SSCs for larger systems that can not be computed in practice can be predicted by extrapolation from SSCs for smaller systems.
- We demonstrated a large potential for accurate TV-based image reconstruction from a reduced number of projections, whereas no potential could be seen for a reduced number of detector bins.
- We showed preliminary indications of a simple relation between gradient image sparsity and the possible undersampling.

6.4.2 Establishing a sparsity-sampling relation

Motivation and goals:

Motivated by the indication of a sparsity-sampling relation we set out to investigate this relation more in detail. While paper A addressed gradient sparsity through TV-minimization, we now focused on sparsity directly in the image domain and reconstruction through 1-norm-minimization. This was to have a more basic problem with better hopes of quantitatively establishing a sparsity-sampling link. Furthermore, while the single image instances indicated a sparsity-sampling relation, we wanted to do a more structured study over multiple similar test images. This prompted us to introduce the concept of image classes and study the sparsity-sampling relation within image classes. By designing image classes with very different structural properties, we obtained a way to simultaneously investigate the influence of structure.

Relevant papers:

[C] Quantitative study of undersampled recoverability for sparse images in computed tomography (page 123).
[G] Connecting image sparsity and sampling in iterative reconstruction for limited angle X-ray CT (page 243).

Methods and results:

In the lack of theoretical results from CS connecting sparsity and sufficient sampling in CT, we carried out an extensive simulation study for determining empirical evidence of a relation. We introduced different classes of phantom images, defined such that randomly generated realizations of image instances from the class could be generated. This was in order to have a well-defined way to generalize single-instance results such as those from subsection 6.4.1.

We generated an ensemble of test images of varying sparsity and reconstructed by 1-norm minimization from data simulating a varying number of projection views in the setting of 2D full-angular range fan-beam CT. For each image instance we recorded the number of projection views required for accurate reconstruction and we studied the ensemble average of the sufficient projection view number. A clear positive relation between the number of nonzeros and the required number of projections views for accurate reconstruction was found. Interestingly, the sparsity-sampling parameter space appeared to be divided into a region of recovery (if data consists of enough views) and nonrecovery (too few views in data), and the transition from one region to the other was very sharp. In other words, images of the same sparsity essentially required the same number of projections to be accurately reconstructed. This was similar to observations made in CS, see, e.g., [48].

We studied this relation under different related set-ups. For example, we considered different image classes with different amounts of spatial structure in the locations of nonzeros. Similar trends were observed but the transition from the nonrecovery region to the recovery region becomes smoother with more structure, i.e., more variation was seen in the required number of projections across image instances.

We studied how the relation scales with the discretization of the image. It turned out that, subject to proper normalization, the same relation could be observed independently of the number of pixels in the image. This means that we can study the sparsity-sampling relation at very finely discretized images through extrapolation of the results at coarser discretization, where it is feasible to run a large number of reconstructions for an ensemble of test images.

We showed that the relation is robust toward small amounts of additive Gaussian noise and discussed how our methodology can be applied to study the relation between sparsity and sufficient sampling for increasingly realistic scenarios as well as in other applications than CT.

We used the same methodology to study the sparsity-sampling relation in the setting of a 2D limited-angle fan-beam CT configuration. A similar relation

was found and the results were used for successfully predicting the number of projection views sufficient for accurate reconstruction in a 3D circular cone-beam CT configuration.

In the general framework of Table 5.1, this work can be considered to belong either to the reconstruction model domain or in the physical domain, depending on the viewpoint taken. If we are interested in determining which one of several reconstruction models that can accurately reconstruct a sparse image from the fewest possible CT projection views, we are in the reconstruction model domain. On the other hand, if we are interested in a single reconstruction model and ask how many projection views are needed for accurate reconstruction, then we are looking at a property of the physical domain (and the imaging model) and, thus, in the physical domain. Such a question could be relevant for choosing between different scanning configurations.

Conclusions:

- We established empirically a quantitative average-case relation between image sparsity and sampling sufficient for accurate reconstruction.
- We argued that the found relation can be used operationally to determine how few projections can be used to reconstruct images with a given expected sparsity.
- We showed existence of a sharp transition between nonrecovery and recovery, i.e., that there is very little variation of the sufficient number of projections for same-sparsity images of same class.
- We showed that the found relations scale linearly with image size, so that the same relation can be expected to exist at more finely discretized images.

The methodologies proposed in these studies can be seen as examples of characterization tools. With the methods of subsection 6.4.1 we can establish full-sampling reference points of DD imaging models to enable a quantitative description of undersampling admitted by sparse reconstruction methods. With a fixed full-sampling reference point, we can compare different reconstruction models. With the methods of subsection 6.4.2 we can study the role of sparsity in a structured way. As such, these studies can be considered blueprints of applying characterization tools. We also note that while both works study relatively small image sizes, we established simple scaling relations so that more realistically sized systems can be studied through extrapolation.

This concludes the description of the main results of the thesis and their connections. In the following chapter, we discuss the results, outline future work and conclude the thesis.

Chapter 7

Discussion and conclusion

The individual thesis contributions are discussed in their specific context in the respective publications enclosed in the appendices. In addition, we give here a general discussion of the contributions in a wider context.

7.1 Discussion and future work

We introduced the general framework of Table 5.1 for describing parameter relations in sparse reconstruction methods in CT. We find that the framework is helpful for analyzing and designing experimental studies, as done in the contributions in chapter 6. However, we realize that adhering to the general framework is a restriction compared to what can actually be done. For example, it may very well be that a heuristic reconstruction method loosely based on for example TV-regularization solved by a non-convergent algorithm produces better images in shorter time, than a method that accurately solves a well-defined problem. However, based only on a reconstructed image, it is difficult to say anything about potential errors committed by such a method. Through studies that do respect the general framework it may be possible to establish relations that can be of use for interpreting the output of such a heuristic reconstruction method.

From an applied perspective, it may be argued that studying only very specific

questions between neighboring domains in the general framework, may take very long time to produce practically applicable reconstruction methods. Moreover, methods can be useful even without a complete understanding of all its parameters, so it may not be worth the effort to carefully characterize every aspect of a reconstruction method. On the other hand, we argue that without an understanding of the underlying relations, it may be possible to obtain single isolated good reconstructions but it will be very hard to reliably ensure good reconstruction on similar cases and to generalize to related cases. Furthermore, once a relation in a particular case has been established it may also be applicable in other cases. In practice, both the practical and the detailed characterization approaches should be used in order to obtain both deep and practical understanding of reconstruction methods.

Our work on algorithms for prototyping reconstruction models may not be considered strictly novel in the mathematical community, because we are working with an existing algorithm and do not propose any improvements to the algorithm itself. However, in addition to simply communicating and making the algorithm accessible to the CT community, our contribution lies in adapting the algorithm and deriving specific "cookbook" algorithm instances to numerous optimization problems of interest to the CT community. It is our perception that the original paper [31], due to its technical nature, would not in itself lead to general adaptation of the proposed algorithm in the CT community. Therefore our intention was to "bridge the gap" by explicitly demonstrating the wide applicability of the algorithm on CT-relevant optimization problems. A natural extension would be to construct a software package with implementation of the prototyping framework, consisting of the generic algorithm and individual "building blocks" to enable the construction of algorithm instances for a variety of optimization problems.

Simplicity, and not efficiency, was the focus for the prototyping algorithms. As long as using the prototyping algorithm only for theoretical studies, where we can wait for the algorithm to converge, this is fine. For practical applications, we need a fast algorithm to avoid excessively long computing time. Ideally, we would have a prototyping algorithm, which in addition to being generally applicable, was extremely fast. The basic prototyping algorithm can be preconditioned for improved efficiency [97]. It would be interesting to study preconditioning techniques specific for CT in order to get closer to a both fast and general prototyping algorithm.

Our work on prototyping algorithms is meant to simplify the study of new reconstruction models, so it is natural in future works to do precisely that. Similar to our paper D in collaboration with SPECT specialists, we should thoroughly investigate new reconstruction models and their design parameters for specific applications.

In our characterization work we obtained empirical evidence of a quantitative relation between image sparsity and the number of CT projection views required for accurate reconstruction. The obtained results focused on sparsity in the image domain, while the more realistic assumption in practice is sparsity in the spatial gradient. It would be interesting to study the relation between gradient sparsity and the number of CT projection views required for accurate reconstruction for example through TV-based reconstruction.

Our characterization studies were limited to simulation studies with synthetic noise-free data in an inverse-crime setting. Natural extensions are to consider more realistic configurations in order to determine how our conclusions generalize. We imagine that a stepwise approach of gradually going to a more realistic configuration will provide the most insight. For example, first replace the inverse-crime simulation study by a non-inverse crime simulation study, then include synthetic noise, then go to actual data of very high-quality and finally address a realistic data set from the application of interest. Such a stepwise approach will reveal the effect of changing each part of the gap between idealized inverse-crime-based conclusions and realistic real-world data.

While our studies were of empirical nature, we find that the results strongly indicate the existence of a relation between sparsity and sufficient sampling. We speculate that an underlying theoretical explanation exists. Perhaps CT system matrices possess a yet unknown property that can explain the observed relation. It would be very interesting to pursue this idea further in an attempt to strengthen the empirical evidence with a theoretical foundation. On the other hand, empirically established relations are often sufficient from an operational perspective. If a relation, such as our observed sparsity-sampling relation, is established empirically, one can be quite confident that the relation can accurately predict the outcome of similar experiment. Empirical evidence, therefore, can directly support in practical experimental design decisions, for example on how many projections to acquire if a given image sparsity is expected.

Our characterization studies focused on the role of the sparsity, but other factors could be studied in a similar structured simulation framework. For example, the tolerable noise level in data and the best choice of regularization parameter. We envision a set of computational characterization tools that an applied researcher can execute on the reconstruction method of interest to provide quantitative understanding of the influence of a range of parameters.

Randomness plays a large role in the currently established results in CS. It would be interesting to determine whether randomized CT sampling, for example through randomly selected projection angles, can lead to theoretical reconstruction guarantees for CT. As an intermediate step, empirical studies, for example using the sparsity-sampling characterization tool of paper C, may be able

to provide useful guidelines. Randomized sampling is not possible in standard scanning configurations of today but if large enough potential is demonstrated in simulation, it may motivate the development of new, randomized scanning configurations.

7.2 Conclusion

In this thesis we have studied sparse image reconstruction methods, as for example proposed in the field of compressed sensing, for the application to CT image reconstruction. Sparse image reconstruction methods can, under suitable circumstances, provide accurate reconstructions from much less data than is required by classical methods. Reconstruction from reduced data is of interest in CT because it corresponds to a reduction of the x-ray exposure, which in turn can lead to reducing the risk of radiation-induced cancer. The use of sparse image reconstruction methods in CT is motivated by the fact that typical cross-section images of the human body consist of fairly well-separated regions of approximately uniform tissue, which leads to sparse spatial gradient images. While the intuitive motivation is clear, we have identified that quantitative understanding of conditions under which sparse image reconstruction methods work well in CT is still lacking. It has been the goal of this thesis to contribute to quantitatively establishing such conditions.

Our main contributions fall in two categories: Algorithm development and empirical characterization of factors affecting the attainable reconstruction quality, in particular the role of the image sparsity in CT.

In the algorithm part, we developed and implemented a number of different algorithms for solving sparse image reconstruction problems, for example the total variation regularization problem. Through simulation studies we investigated various algorithmic aspects including convergence rate dependence on problem parameters and the choice of termination criterion. We developed a reconstruction model prototyping framework to enable applied researchers to effortlessly obtain convergent optimization algorithms for a host of optimization problems of interest to CT.

In the characterization part, we focused primarily on studying the influence of image sparsity. We established quantitatively an average-case relation between the image sparsity and the possible reduction in data that allows for accurate reconstruction. We found similar relations to hold across different classes of images and, thus, our results provide applied researchers with a quantitative way of determining how large a degree of undersampling (or reduction in x-ray

exposure) that can be applied for reconstruction of images of a given sparsity.

The specific results in the thesis work are all obtained in the setting of medical CT. However, the algorithms developed and the general characterization methodology are not limited to the case of medical CT. In contrast, we foresee that significant fundamental insight can be obtained by use of the proposed tools and generalizations hereof in other applications of tomographic imaging.

Communication and collaboration between the CT community and the mathematical imaging community is very limited today, possibly due to the necessity of thoroughly understanding the other field's more practical or theoretical aspects, respectively. We find that increased levels of communication and interest in collaboration will be necessary if sparse reconstruction methods are to become of practical daily use in CT. The present thesis is a step in this direction.

Appendix A

Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in x-ray CT

IEEE Trans. Med. Imaging, vol. 32, issue 2, pp. 460–473, 2013.
doi:10.1109/TMI.2012.2230185.
First published online 27 November 2012, in print 30 January 2013.

J. S. Jørgensen, E. Y. Sidky and X. Pan

460 IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 32, NO. 2, FEBRUARY 2013

Quantifying Admissible Undersampling for Sparsity-Exploiting Iterative Image Reconstruction in X-Ray CT

Jakob S. Jørgensen*, Emil Y. Sidky, and Xiaochuan Pan

***Abstract*—Iterative image reconstruction with sparsity-exploiting methods, such as total variation (TV) minimization, investigated in compressive sensing claim potentially large reductions in sampling requirements. Quantifying this claim for computed tomography (CT) is nontrivial, because both full sampling in the discrete-to-discrete imaging model and the reduction in sampling admitted by sparsity-exploiting methods are ill-defined. The present article proposes definitions of full sampling by introducing four sufficient-sampling conditions (SSCs). The SSCs are based on the condition number of the system matrix of a linear imaging model and address invertibility and stability. In the example application of breast CT, the SSCs are used as reference points of full sampling for quantifying the undersampling admitted by reconstruction through TV-minimization. In numerical simulations, factors affecting admissible undersampling are studied. Differences between few-view and few-detector bin reconstruction as well as a relation between object sparsity and admitted undersampling are quantified.**

***Index Terms*—Compressed sensing (CS), computed tomography (CT), data models, image sampling, iterative methods.**

I. Introduction

RECENTLY, iterative image reconstruction (IIR) algorithms have been developed for X-ray tomography [1]–[10] based on the ideas discussed in the field of compressive sensing (CS) [11]–[14]. These algorithms promise accurate reconstruction from less data than is required by standard image reconstruction methods. This is made possible by exploiting sparsity, i.e., few nonzeroes in the image or of some transform applied to the image. One can argue about whether these algorithms are truly novel or not: edge-preserving regularization and reconstruction based on the total variation (TV) semi-norm [15]–[18] have a clear link to sparsity in the object gradient and have been considered before the advent of CS, and algorithms specifically for object sparsity have been developed for blood vessel imaging with contrast agents [19]. Nevertheless, the interest in CS has broadened the perspective on applying optimization-based methods for IIR algorithm development for computed tomography (CT), and it has motivated development of efficient algorithms involving variants of the ℓ_1-norm [10], [20]–[25].

What is seldom discussed, however, is that the theoretical results from CS do not extend to the CT setting. CS only provides theoretical guarantees of accurate undersampled recovery for certain classes of random measurement matrices [13], not deterministic matrices such as CT system matrices. While the mentioned references demonstrate empirically that CS-inspired methods do indeed allow for undersampled CT reconstruction, there is a fundamental lack of understanding of why, and the conditions under which, this is the case. One problem in uncritically applying sparsity-exploiting methods to CT is that there is no quantitative notion of full sampling.

Most IIR, including sparsity-exploiting, methods employ a discrete-to-discrete (DD) imaging model[1] which requires that the object function be represented by a finite-sized expansion set and sampling specified over a finite set of transmission rays. This contrasts with most analysis of CT sampling, which is performed on a continuous-to-discrete (CD) model. For analyzing analytic algorithms such as filtered back-projection (FBP), a continuous-to-continuous (CC) model such as the X-ray or Radon transform is chosen, and discretization of the data space is considered, yielding results for the corresponding CD model. Analysis of the CD model is performed independent of object expansion. If the expansion set for the DD model is chosen to be point-like, e.g., pixels/voxels, there may be similarity between CD and DD models justifying some crossover of intuition on sampling, but in general sufficient-sampling conditions can be different for the two models. That a more fine-grained notion of sufficient sampling is needed for the DD model can be seen by considering the representation of the object function on a 128×128 versus a 1024×1024 pixel array. Clearly, the latter case requires more samples than the former, but sampling conditions derived from the CD model cannot make this distinction. Sufficient sampling for the DD model becomes even less intuitive for nonpoint-like expansion sets such as natural pixels, wavelets, or harmonic expansions. Yet, to quantify the level of *under*sampling admitted by a sparsity-exploiting IIR method, *full* sampling needs to be defined

Manuscript received August 16, 2012; revised October 25, 2012; accepted November 17, 2012. Date of publication November 27, 2012; date of current version January 30, 2013. This work was supported in part by the project CSI: Computational Science in Imaging under Grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences, and in part by the National Institutes of Health (NIH) R01 under Grant CA158446, Grant CA120540, and Grant EB000225. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health. *Asterisk indicates corresponding author.*

*J. S. Jørgensen is with the Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Lyngby, Denmark (e-mail: jakj@imm.dtu.dk).

E. Y. Sidky and X. Pan are with the Department of Radiology, University of Chicago, Chicago, IL 60637 USA (e-mail: sidky@uchicago.edu; xpan@uchicago.edu).

Digital Object Identifier 10.1109/TMI.2012.2230185

[1]See [26, Ch. 15] for an overview of different imaging models.

for the corresponding DD model, and to that end we introduce several sufficient-sampling conditions (SSCs).

Specifically, in the present article, SSCs for the DD model are derived from the condition number of the corresponding system matrix. Multiple SSCs are defined to characterize both invertibility and stability of the system matrix. To perform the analysis, a class of system matrices is defined so that the system matrix depends on few parameters. The class is chosen so that it has wide enough applicability to cover thoroughly a configuration/expansion combination of interest, but not so wide as to make the analysis impractical. For the present study, we select a system matrix class for a 2-D circular fan-beam geometry using a square-pixel array. The SSCs are chosen so that they provide a useful characterization of any system matrix class, but the particular values associated with the SSCs in this work apply only to the narrow system matrix class defined. While the article presents a strategy for defining full sampling, the analysis must be redone with any alteration to the system matrix class.

After deriving the SSCs for the particular circular fan-beam CT system matrix class, we apply sparsity-exploiting IIR in the form of constrained TV-minimization. We consider the specific application of CT to breast imaging and use a realistic and challenging discrete phantom. We use the SSCs as reference points of full sampling for quantifying the undersampling admitted by each of the conducted reconstructions. Specifically, we demonstrate significant differences in undersampling admitted for reconstruction from few views compared to few bins. We study how variations to the reconstruction optimization problem, to the image quality metric, to the discretization method for the system matrix, and to the sparsity of the phantom image affect the results.

In Section II we describe the CT imaging model and present the particular system matrix class we employ for circular fan-beam CT reconstruction. In Section III we give a background on sparsity-exploiting methods. In Section IV the SSCs are presented and their application is illustrated for the 2-D circular fan-beam case. Finally, Section V illustrates an example study on quantifying admissible undersampling by constrained TV-minimization employing the discussed SSCs.

II. Class of System Matrices for the Discrete-to-Discrete Imaging Model

A. The X-Ray Transform

Explicit image reconstruction algorithms such as FBP are based on inversion formulas for the CC cone-beam or X-ray transform model

$$g[\vec{s},\vec{\theta}] = \int_0^\infty dt\, f(\vec{s}+t\vec{\theta}) \tag{1}$$

where g, the line integral over the object function f from source location $\vec{s}$ in the direction $\vec{\theta}$, is considered data. Fan-beam FBP, for example, inverts this model for the case where the source location $\vec{s}$ varies continuously on a circular trajectory surrounding the subject, and at each $\vec{s}$ the ray-direction $\vec{\theta}$ is varied continuously through the object in the plane of the source trajectory.

B. The Discrete-to-Discrete Model

For most IIR algorithms, the CC imaging model is discretized by expanding the object function in a finite expansion set, for example, in pixels/voxels. Furthermore, the discrete digital sampling of the CT device is accounted for by directly using the sampled data without interpolation. The effect of both of these steps is to convert the imaging model to a discrete-to-discrete (DD) formulation

$$\vec{g} = X\vec{f} \tag{2}$$

where $\vec{g}$ represents a finite set of ray-integration samples, $\vec{f}$ are coefficients of the object expansion, and X is the system matrix modeling ray integration. This DD imaging model is almost always solved implicitly, because the matrix X, even though sparse, is beyond large for CT applications: X is in the domain of a giga-matrix for 2-D imaging and a tera-matrix for 3-D imaging.

A central point motivating the strategy of the present work is that the DD imaging model has a more focused scope than the CD model, because the former can often be derived from the CD model by expanding the continuous image domain with a finite set of functions. How the discretization of the CD model is done for CT to achieve the DD imaging model is not standardized. Many expansion elements have been used in CT studies; in addition to pixels, for example blobs [27], wavelets [28], [29], and natural pixels [30], [31]. Also, the matrix elements using only the pixel expansion set can be calculated in different ways that all tend toward the CC model in the limit of shrinking pixel size and detector bin size. Different modeling choices will necessarily alter X. This tremendous variation in X means that it is important to fully specify X for each study, and it is important to recharacterize X for any change in the model. For example, changing pixel size can have large impact on the null space of the system matrix in the DD model.

In order to describe precisely and provide a delimitation of the system matrices considered in the present work, we introduce the notion of a *system matrix class*. Any given system matrix depends on numerous model parameters determining the scanning geometry, sampling and discrete expansion set. A system matrix class consists of the system matrices arising from fixing a number of these parameters and leaving a subset of the parameters free. The system matrix class can then be studied by varying these free parameters.

C. The System Matrix Class Used in the Present Study

In CT, projections are acquired from multiple source locations which lie on a curve trajectory and the source location $\vec{s}(\lambda)$ is specified by the scalar parameter λ. The circular trajectory is the most common, and is what we use here

$$\vec{s}(\lambda) = R_0(\cos\lambda, \sin\lambda)$$

where R_0 is the distance from the center-of-rotation to the X-ray source, and set to $R_0 = 40\ \mathrm{cm}$ in the present work. The detector bin locations are given by

$$\vec{b}(\lambda, u) = (R_0 - D)(\cos\lambda, \sin\lambda) + u(-\sin\lambda, \cos\lambda)$$

where D is the source-to-detector-center distance ($D = 80$ cm in the present work), and u specifies a position on the detector. The ray direction for the detector-geometry independent data function is

$$\vec{\theta}(\lambda, u) = \frac{\vec{b}(\lambda, u) - \vec{s}(\lambda)}{\|\vec{b}(\lambda, u) - \vec{s}(\lambda)\|_2}.$$

The 2π arc is divided into N_{views} equally spaced angular intervals, so that the source parameters follow:

$$\lambda_i = i\Delta\lambda \tag{3}$$

where

$$\Delta\lambda = \frac{2\pi}{N_{\text{views}}} \text{ and } i \in [0, N_{\text{views}} - 1]. \tag{4}$$

The detector is subdivided into N_{bins}

$$u_j = u_{\min} + (j + 0.5)\Delta u \tag{5}$$

where D_L is the detector length ($D_L = 41.3$ cm), $u_{\min} = -D_L/2$, $\Delta u = D_L/N_{\text{bins}}$, and $j \in [0, N_{\text{bins}} - 1]$. The detector length is determined by requiring it to detect all rays passing through the largest circle inscribed within the square $N \times N$ image array for which we use the side length 20 cm. We restrict the unknown pixel values to lie within this circular field-of-view (FOV), and the number of unknown pixel values N_{pix} is

$$N_{\text{pix}} \approx \frac{\pi}{4} N^2 \tag{6}$$

where the actual value, which has to be an integer, is given with each simulation below.

Effectively, the dimensions of the projector X are $M = N_{\text{views}} \times N_{\text{bins}}$ rows (number of ray integrations) and N_{pix} columns (number of variable pixels). To obtain the individual matrix elements, the line-intersection method is employed, where $X_{m,n}$ is the intersection length of the mth ray with the nth pixel. This description completely specifies the system matrix class for the present circular fan-beam CT study, and the free parameters of this class are N, N_{views}, and N_{bins}.

III. CT Image Reconstruction by Exploiting Gradient-Magnitude Sparsity

Reconstruction of objects from undersampled data within the DD imaging model corresponds to a measurement matrix X with fewer rows than columns. The infinitely many solutions are narrowed down by selecting the sparsest one, i.e., the one that has the fewest number of nonzeroes, either in the image itself or after some transform has been applied to it. Mathematically, the reconstruction can be written as the solution of the constrained optimization problem

$$\vec{f}^* = \arg\min_{\vec{f}} \|\Psi(\vec{f})\|_0 \text{ such that } X\vec{f} = \vec{g}. \tag{7}$$

Here, Ψ is a sparsifying transform, for instance a discrete wavelet transform, and $\|\cdot\|_0$ is the ℓ_0-"norm" (although it is in fact not a norm), which computes its argument vector's sparsity, that is, counts the number of nonzeroes. The equality constraint restricts image candidates to those agreeing exactly with the data.

Central results in CS derive conditions on X drawn from certain random system matrix classes such that $\vec{f}^*$ is exactly equal to the underlying unknown image that gave rise to the data $\vec{g}$. Two key elements are sparsity of $\Psi(\vec{f})$ and incoherence of X: exact recovery depends on the size of $\vec{g}$ being larger than some small factor of $\|\Psi(\vec{f})\|_0$ [13], and the concept of incoherence is needed to ensure that the few measurements $\vec{g}$ available give meaningful information about the nonzero elements of $\Psi(\vec{f})$. Other important results in CS involve the relaxation of the nonconvex ℓ_0-"norm" to the convex ℓ_1-norm

$$\vec{f}^* = \arg\min_{\vec{f}} \|\Psi(\vec{f})\|_1 \text{ such that } X\vec{f} = \vec{g}. \tag{8}$$

In contrast to (7), this convex problem is amenable to solution by a variety of practical algorithms, although the large scale of CT matrices still presents a challenge for algorithm development. Another important contribution from CS is the derivation of conditions under which the solution to (8) is identical to the solution to (7), so that the sparsest solution can be found by solving (8).

For application to medical imaging, it was suggested in [12] that a potentially useful Ψ would be to have Ψ compute the discrete gradient magnitude of $\vec{f}$, i.e., for the jth pixel

$$\left[\Psi(\vec{f})\right]_j = \|D_j\vec{f}\|_2 \tag{9}$$

where D_j computes a finite-difference approximation of the gradient at each pixel j, and the two-norm also acts pixel-wise on the differences. In CT, for example, the typical image consists of regions having an approximately constant gray-level value separated by sharp boundaries between various tissue types. The magnitude of the spatial gradient of such images is zero within constant regions and nonzero along edges, so the gradient magnitude image can be sparse. The ℓ_1-norm applied to the gradient magnitude image is known as the total variation (TV) semi-norm

$$\|\vec{f}\|_{\text{TV}} = \|\Psi(\vec{f})\|_1 = \sum_{j=1}^{N_{\text{pix}}} \|D_j\vec{f}\|_2 \tag{10}$$

and the optimization problem of interest becomes

$$\vec{f}^* = \arg\min_{\vec{f}} \|\vec{f}\|_{\text{TV}} \text{ such that } X\vec{f} = \vec{g}. \tag{11}$$

However, the theoretical results from CS do not extend to the CT setting. Three properties that separate CT matrices from typical CS matrices are that CT matrices:

1) are structured and do not belong to random matrix classes for which CS results are proved [13];
2) can have rank smaller than the number of rows, which means that there exist vectors $\vec{g}$ in the data space that are inconsistent with X, and accordingly the linear imaging model (2) has no solution;
3) may be numerically ill-conditioned in case of having more rows than columns (data set size is greater than the image representation).

Nevertheless, it has been demonstrated empirically in extensive numerical studies with computer phantoms under ideal data conditions as well as with actual scanner data that highly accurate reconstructed images for "undersampled" projection data can be obtained from (8) and variants thereof.

It is precisely this last phrase which is of interest in the present paper: what exactly does it mean to have "undersampled" data for CT? Undersampled data implicitly relies on a certain level of sampling being sufficient—but no such precise concept exists for CT using the DD imaging model, to the best of our knowledge. Without a reference point for having sufficient sampling it is difficult to quantify admissible levels of undersampling. In the present paper, we aim to provide this reference point. Specifically in Section IV, we propose sufficient-sampling conditions (SSCs) to be computed for specific system matrix classes, and which serve as a reference for quantifying the admissible undersampling for sparsity-exploiting reconstruction. Application of the SSCs is demonstrated with numerical simulations of breast CT.

IV. Sufficient-Sampling Conditions

In considering sufficient sampling for circular fan-beam CT, the CC model is recast as a CD model by introducing a discrete sampling operator, usually taken to be evenly distributed delta functions, on the CT sinogram space. Making the assumption that the underlying sinogram function is band-limited, many useful and widely applicable results have been obtained, see for example Sec. 3.3 of [32] and [33]–[35]. Furthermore, for more advanced scanning geometries and sampling patterns there are available tools for analysis such as singular value decomposition (SVD), direct analysis of multi-dimensional aliasing, and the evaluation of the Fourier crosstalk matrix [36]–[38]. These important results, however, do not apply directly to IIR, since for the DD model we need to take into account the finite image expansion set.

We consider an empirical approach for characterizing sufficient sampling within a class of system matrices. The idea is to fix the image representation, which for the circular fan-beam system matrix class is the parameter N_{pix}, and then vary the sampling parameters N_{views} and N_{bins} to ensure accurate determination of the pixel values. This is done by establishing sufficient-sampling conditions (SSCs) based on matrix properties in the considered system matrix class. If the system matrix class is altered, the SSC-analysis must be redone.

A. SSC Definitions

The SSCs, we propose, characterize invertibility and ill-conditioning of the system matrix class. In considering the DD imaging model (2), the data $\vec{g}$ are restricted to the range of X. This separates out the issue of model inconsistency which does not have direct bearing on sampling conditions.

We define SSC1 to be sampling such that X has at least as many rows as columns. If there are fewer rows than columns, X necessarily has a nontrivial null space, and solutions of (2) will not be unique. Even if the number of rows is equal to or larger than the number of columns, there may still be a nontrivial null space, because the rows can be linearly dependent. In addition to SSC1, we define SSC2 to mean that X has a null space consisting only of the zero image, or equivalently, that the smallest singular value $\sigma_{\min}$ of X is nonzero. Existence of a unique solution to (2) is ensured by SSC2. Both SSC1 and SSC2 can be evaluated for any system matrix class.

Neither of SSC1 and SSC2 address numerical instability and to address that, we employ the condition number of X, the ratio of the largest and smallest singular values

$$\kappa(X) = \frac{\sigma_{\max}}{\sigma_{\min}}. \tag{12}$$

The condition number κ can be as small as 1 and the larger κ becomes, the more numerically unstable is solution of $X\vec{f} = \vec{g}$. How to use κ to define a SSC requires some discussion.

Whereas sensing matrix classes studied in CS typically are well-conditioned—for instance the square discrete Fourier transform (DFT) matrix is orthogonal, thus having a condition number of 1—the system matrices encountered in X-ray CT can have a relatively large condition number [32], [39], which leads to numerical instability and thus large sensitivity to noisy measurements. Even if SSC2 holds, the condition number κ is finite but may still be large, potentially allowing other images than the desired solution to be numerically close to satisfying $X\vec{f} = \vec{g}$. If we fix the image representation, which for the present 2-D circular fan-beam setup amounts to fixing N_{pix}, and increase the sampling, allowing N_{views} and N_{bins} to increase toward ∞, the condition number will decrease toward a limiting condition number

$$\kappa_{\mathrm{DC}} = \lim_{N_{\mathrm{views}}, N_{\mathrm{bins}} \to \infty} \kappa(X) \tag{13}$$

where the DC subscript refers to the fact that X is limiting to a discrete-to-continuous (DC) system matrix[2]. The limiting condition number κ_{DC} is the best-case κ for a fixed image representation, but κ_{DC} may still be larger than 1.

For actual CT scanners, it is not practical to allow N_{views} and N_{bins} to increase without bound, and empirical experience shows diminishing improvements in doing this. To balance the impracticality of going to continuous sampling on the one hand against the need to optimize numerical stability on the other, we introduce SSC3 to mean that the condition number of X satisfies

$$\frac{\kappa(X)}{\kappa_{\mathrm{DC}}} < r_{\mathrm{samp}} \tag{14}$$

where r_{samp} is a finite ratio parameter greater than 1. The smaller the choice of r_{samp}, the closer X is to the DC limit. This SSC can also be generally applied to other system matrix classes, but the appropriate parameter setting of r_{samp} will be specific to a particular class.

Finally, we introduce SSC4 specifically for the present 2-D circular fan-beam system matrix class. This SSC is taken to mean $2N$ samples in both the view and bin directions, i.e., $N_{\mathrm{views}} = N_{\mathrm{bins}} = 2N$. This SSC is simple to evaluate, and we will demonstrate empirically that it is a useful condition, which acts as a good approximation for attaining SSC3 with $r_{\mathrm{samp}} = 1.5$. This SSC is specific to the system matrix class

[2]We conjecture that the limiting condition number is finite and well-defined based on empirically observing convergence to a constant with increasing N_{views} and N_{bins} for the considered system matrix class.

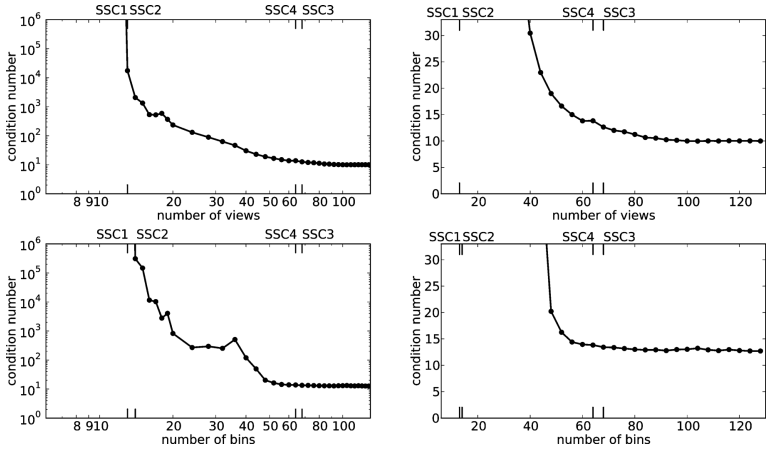


Fig. 1. Condition numbers for system matrices (line-intersection) modeling circular fan-beam projection data from the 812-pixel circular FOV contained within a 32×32 pixel square. Top: number of bins is fixed at 64. Bottom: the number of views is fixed at 64. Left: Double-logarithmic plot for overview. Right: Linear plot of details. The abbreviations SSC1, SSC2, SSC3, and SSC4 are the sufficient-sampling conditions discussed in Section IV-A.

investigated here. Even slightly different system matrix classes might not allow for the same SSC4 definition.

Our strategy is similar to analysis presented in early works on CT, such as in [39], but the point here is not novelty of the analysis; rather we need to establish a reference point by which to evaluate the sampling reduction admitted by sparsity-exploiting methods.

In what follows, the proposed SSCs are examined for the 2-D circular fan-beam system matrix class. First, small systems are considered, where X can be explicitly computed and analyzed so that the full set of singular values of X is attainable. Second, we argue that our conclusions generalize to larger, more realistic systems, where X is impractical to store in computer memory and it is only feasible to compute the smallest and largest singular values.

B. SSCs for Small Systems

We consider a small $N = 32$ image array with $N_{\text{pix}} = 812$, and generate system matrices X for different numbers of views, $N_{\text{views}} \in [8, 128]$, and detector bins, $N_{\text{bins}} \in [8, 128]$. The condition number $\kappa(X)$ is computed through direct SVD of X for all values of N_{views} and N_{bins} within the specified parameter ranges, and $\kappa(X)$ as a function of N_{views} (for fixed $N_{\text{bins}} = 64$) and N_{bins} (for fixed $N_{\text{views}} = 64$) is shown in Fig. 1. The plots show three phases: the left-most part, where the condition number is infinite; the middle, where the condition number becomes finite and decays slowly; and the right-most part, where it remains relatively stable. The positions of the different SSCs are shown at the top and bottom, and they serve as transition points between the three phases.

For varying the number of views, SSC1 occurs at $N_{\text{views}} = 13$, where X is of size 832×812. In fact, also SSC2 occurs here, since κ has become finite. For varying the number of bins, SSC1 occurs at the same place, but SSC2 needs 14 bins. In general, we have no way to determine whether SSC1 and SSC2 occur in the same position for the whole system matrix class, which makes SSC1 less reliable as a general reference of full sampling. On the other hand, SSC2 is a reliable reference point for full sampling, however, SSC2 requires more work to determine, because its location can change with a change of system matrix class.

After passing SSC2, the condition number κ decreases. For larger N_{views} and N_{bins}, the decay becomes slower, and we pick $r_{\text{samp}} = 1.5$ as a trade-off between a sufficiently small condition number and a finite number of views. As an approximation of κ_{DC} we take the value of κ at $N_{\text{views}} = N_{\text{bins}} = 4N = 128$, yielding $\kappa_{\text{DC}} = 9.17$. Then SSC3 occurs at $N_{\text{views}} = 68$ and at $N_{\text{bins}} = 68$, which suggests a symmetry in N_{views} and N_{bins}. On the other hand, the decrease in κ during the middle part is not symmetric in the parameters N_{views} and N_{bins}; the decrease in κ with N_{views} is gradual while that of N_{bins} is step-like at $N_{\text{bins}} = 48$. Nevertheless, at the position of SSC3, there is only small further reductions in κ to be gained by going to larger N_{views} and N_{bins}. The simpler condition SSC4 occurs at $N_{\text{views}} = 64$ and $N_{\text{bins}} = 64$, and it approximates SSC3 with $r_{\text{samp}} = 1.5$ closely.

C. Altering the System Matrix Class

Altering the system matrix class will in general alter the SSCs. To demonstrate this effect, we replace the line-intersection based system matrix class by ray-tracing, using nearest-neighbor interpolation at the mid-line of each pixel row. The experiment is repeated and the obtained condition numbers are shown in Fig. 2, along with the ones based on line-intersection, for comparison. The shown SSCs are for

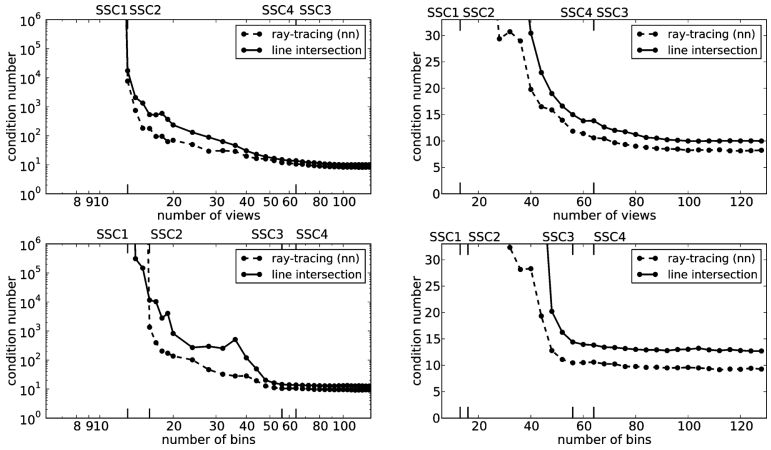


Fig. 2. Condition numbers for system matrices (ray-tracing with nearest neighbor interpolation, and line-intersection for comparison) modeling circular fan-beam projection data from the 812 pixels circular FOV contained within a 32×32 pixel square. Top: number of bins is fixed at 64. Bottom: the number of views is fixed at 64. Left: double-logarithmic plot for overview. Right: linear plot of details. The labels SSC1, SSC2, SSC3, and SSC4 are the sufficient-sampling conditions, discussed in Section IV-A, for the ray-tracing system matrix.

ray-tracing. While the same overall trends are seen, there are some significant differences. First, for fixed $N_{\text{views}} = 64$ we need $N_{\text{bins}} = 16$ to obtain SSC2, compared to 14 for line-intersection. This firmly establishes that SSC1 does not imply SSC2, and that the precise position of SSC2 cannot be inferred from knowing SSC2 of a similar system matrix. Second, the ray-tracing condition numbers are smaller than the same for line-intersection, for instance, $k_{\text{DC}} = 7.23$ is 20% lower relative to the line-intersection version of X. That the ray-tracing condition numbers are lower does not necessarily mean that this method is "better" than the line-intersection method for real-world applications, because the other side of the story is model error, which is not considered here. Finally, the positions of SSC3 are different, for fixed $N_{\text{bins}} = 64$ coinciding with SSC4, while for fixed $N_{\text{views}} = 64$ occurring at $N_{\text{bins}} = 56$. Still, for larger N_{views} and N_{bins} there are only small further reductions in κ to be gained, and SSC4, at $r_{\text{samp}} = 1.45$, approximates SSC3 closely.

One could imagine that other system matrix classes such as employing area-weighted integration instead of the linear integration or different basis functions could alter the condition numbers of X even more substantially. We do not include results for more system matrix classes, as our goal here is not to provide a comprehensive comparison between all conceivable classes, but merely to stress that different classes can have different SSCs, and to propose carrying out the same study for gaining insight in the particular system matrix class at hand.

D. SSCs for Larger Systems

The results shown in Figs. 1 and 2 give a sense about various sampling combinations, but the system size is unrealistically small. In this section we aim to extend the results to larger systems. For large X, it is not practical to compute the direct SVD for evaluating κ. Instead, we seek only to obtain $\sigma_{\min}$ and $\sigma_{\max}$, which can be accomplished through the power and inverse power methods [40].

For characterizing the present circular, fan-beam system matrix class, $\kappa(X)$ is computed for larger image arrays with sizes $N = 32, 64, 128, 256$. We focus on sampling conditions where $N_{\text{bins}} = 2N$ and report the data sampling in views, N_{views}, as multiples of N, ranging from 1.0 to 4.0. The left plot in Fig. 3 shows the condition number as a function of N for each sampling size on a double logarithmic scale. A clear linear trend is seen in all cases and the best linear fits and their slopes are also shown, in all cases very close to 0.50, and we conclude that κ scales with $\sqrt{N}$. For increasing N_{views}, the condition numbers tend towards the bottom line, note in particular that not much difference is seen between $N_{\text{views}} = 3N$ and $N_{\text{views}} = 4N$ indicating that the limiting κ is approached. We conclude that κ_{DC} also scales with $\sqrt{N}$. As a result, it can expected that at SSC4, i.e., $N_{\text{views}} = N_{\text{bins}} = 2N$, $\kappa(X)/\kappa_{\text{DC}} \approx 1.5$, which was the case at $N = 32$. Hence, SSC4 will continue to approximate SSC3 closely, when the image size is increased. To further support this conclusion we show in the right plot of Fig. 3 the ratio $r_{\text{samp}} = \kappa(X)/\kappa_{\text{DC}}$ as function of the number of views (normalized by N) for each N. The r_{samp} values are almost identical for all N and intersect the line $r_{\text{samp}} = 1.5$ very close to $N_{\text{views}} = 2N$, which is precisely SSC4.

E. Summary of SSCs

The conditions SSC1 and SSC2 are useful reference points for invertibility of X and can be computed for any system matrix

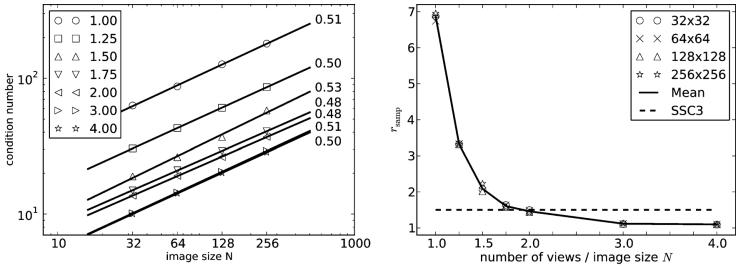


Fig. 3. Left: condition numbers as function of image size. Each symbol represents different number of views ranging from N to $4N$. Circular image arrays are used with sizes given by $N = 32, 64, 128, 256$. The number of bins is fixed at $2N$. With each number of views is also shown the best linear fit and its slope is given. In all cases the condition number scales with $\sqrt{N}$. Right: same condition numbers normalized by the respective κ_{DC} at each N and plotted as function of view number normalized by image size N. The full line is the point-wise mean and the dashed line is the position of SSC3 with $r_{\mathrm{samp}} = 1.5$. Independently of N, SSC3 with $r_{\mathrm{samp}} = 1.5$ occurs very close to $N_{\mathrm{views}} = 2N$.

class. The size of the gap between SSC1, X being square, and SSC2, X having an empty null space, is governed by inherent linear dependence of the rows of the system matrix. Because the results show little difference between SSC1 and SSC2 for the present system matrix class and SSC1 is easier to compute, we use only SSC1 in the simulation studies in Section V.

For system matrix classes representing CT imaging, stability of the system matrix plays an important role, and accordingly we have introduced SSC3 which also can be computed for any system matrix class. For the present, circular fan-beam system matrix class, SSC3 at $r_{\mathrm{samp}} = 1.5$ is a useful operating point, and this level of sampling is well approximated by the simple rule, SSC4, where $N_{\mathrm{views}} = N_{\mathrm{bins}} = 2N$. We point out that for other system matrix classes, even those representing circular fan-beam CT, other operating points for SSC3 may be more appropriate and empirical studies must be performed to see if a simple condition, such as SSC4, can approximate accurately SSC3.

In the remaining part of the paper we demonstrate how we can use the SSCs as a reference for stating admissible undersampling factors in sparsity-exploiting reconstruction.

V. Numerical Experiments With SSCs and Sparsity-Exploiting Undersampled Reconstruction

In this section we investigate sparsity-exploiting IIR in numerical simulation studies. Our goal is to numerically demonstrate and quantify the undersampling admitted by sparsity-exploiting IIR, i.e., at which an accurate reconstruction is obtained. We use numerical simulation, because we are unaware of any theoretical results establishing undersampling guarantees for the present system matrix class. We focus here on exploiting gradient-magnitude sparsity by use of constrained TV-minimization.

Three important factors differentiate the present studies from previous simulation work with constrained TV-minimization:

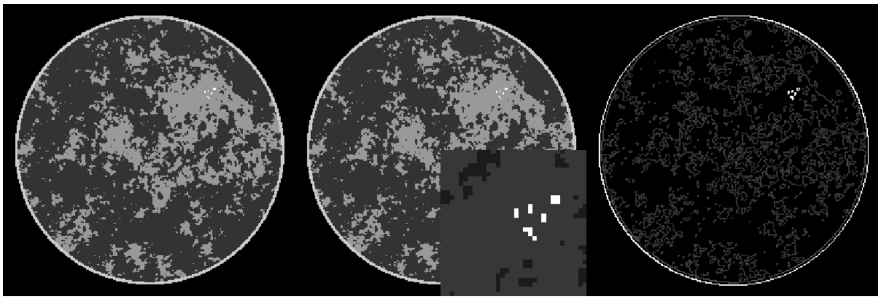
1) use of phantoms with realistic complexity;
2) numerically accurate solution to the constrained TV-minimization problem;
3) quantitative references for full sampling—the central topic of the paper.

For each factor, we briefly discuss the significance.

Much simulation work on constrained TV-minimization has used regular, piece-wise constant phantoms, such as the Shepp–Logan phantom, to demonstrate the promise of the technique. For that purpose, such unrealistically simple phantoms were fine, and simulations were generally followed up by demonstration with actual CT projection data. For the present purpose of quantifying admissible undersampling, we need phantoms with similar complexity as would be encountered in CT applications, and as an example we focus on breast CT. The standard measure of complexity employed in CS is the image sparsity, i.e., the number of nonzeroes in the image, or in the case of TV-minimization the number of nonzeroes in the gradient-magnitude image. Accordingly, we choose a digital phantom with realistic gradient-magnitude sparsity modeling breast anatomic structure [41].

The accuracy requirement on the solver of constrained TV-minimization for the present study is extremely high. The optimization problems in Section V-B, below, are solved to high accuracy, which has been made possible only recently for large-scale CT problems involving the TV-semi-norm through development of advanced first-order methods [21], [22], [25]. This level of accuracy is necessary, because empirical image error results obtained by sweeping parameters of the system matrix class will be used for determining whether a numerically computed solution is close to the original. High-accuracy solutions remove any doubt about whether the resulting images, and the corresponding quantitative measures, depend on the algorithm used to solve constrained TV-minimization.

The SSCs defined above provide reference points useful for interpreting the empirical results of this section and help to quantify undersampling admitted by constrained TV-minimization.



(a) (b) (c)

Fig. 4. Left: 256×256 pixelized breast CT phantom used in the present study. Middle: same with ROI around microcalcifications shown magnified as inset. Right: the gradient magnitude image, which has a sparsity of 10019 nonzero pixel values.

A. Breast CT Background

Breast CT [42], [43] is being considered as a possible screening or diagnostic tool for breast cancer. The system requirements are challenging from an engineering standpoint, because this type of CT must operate with a total exposure similar to two full-field digital mammograms (FFDM). FFDMs for a screening exam entail two X-ray projections, while breast CT acquires on the order of 500 X-ray projections. The exposure previously used for only two views is now divided up among 250 times more projections. Accordingly, sparsity-exploiting IIR algorithms for CT may have an impact on the breast CT application. The potential to reconstruct volumes from fewer views than a typical CT scan might allow an increased exposure per view.

For the present study, we employ the breast phantom originally described in [41] and displayed in Fig. 4. It consists of $N_{\mathrm{pix}} = 51\,468$ pixels within the circular image region, contained in a 256×256 array. The breast phantom has a small region of interest (ROI) containing five tiny ellipses which model microcalcifications. The gray values range from 1.0 to 2.3, normalized to the background fat tissue, hence unit-less. The modeled tissues and corresponding gray values are fat at 1.0, fibroglandular tissue at 1.10, skin at 1.15, and microcalcifications ranging from 1.9 to 2.3. The sparsity in the gradient magnitude image is 10019, or roughly one fifth of N_{pix}. Because we are investigating the utility of gradient-magnitude sparsity-exploiting algorithms, it is important that the test phantom have a realistic sparsity level relative to the actual application.

B. Simulation Optimization Problems and Algorithms

Our goal is to evaluate quantitatively what level of undersampling reconstruction through (11) allows. Similar to the analysis of the linear imaging model (2), only data $\vec{g}$ in the range of X is considered. Although not a realistic assumption for actual CT data, this "inverse crime" scenario [44] is appropriate for obtaining a reference of the underlying admissible undersampling. For the numerical studies, we solve a relaxed form of (11), where the data equality constraint is replaced by an inequality allowing for a small deviation ϵ from data as measured by the distance D between the data $\vec{g}$ and the projection $X\vec{f}$ of some image $\vec{f}$

$$D(X\vec{f}, \vec{g}) \leq \epsilon \tag{15}$$

where

$$D(X\vec{f}, \vec{g})^2 = \frac{1}{N_{\mathrm{views}} N_{\mathrm{bins}}} \|X\vec{f} - \vec{g}\|_2^2.$$

Scaling the data error D with N_{views} and N_{bins} is done to enable comparison across images reconstructed from different view and detector bin numbers. The constrained TV-minimization problem is

$$\ell_2 - \mathrm{TV}: \quad \vec{f}^{\,*} = \arg\min_{\vec{f}} \|\vec{f}\|_{\mathrm{TV}}. \tag{16}$$
$$\text{subject to} \quad D(X\vec{f}, \vec{g}) \leq \epsilon.$$

Accurate solution of (16) is nontrivial; although the objective is convex, it is not quadratic. The algorithm employed here solves its Lagrangian using an accelerated first-order method, using only the objective and its gradient, and is explained in detail in [21]. An important technical detail for this algorithm is that it requires that the image TV-term be differentiable. For the algorithm implementation we use a smoothed TV-term, $\sum_j \sqrt{\|D_j\vec{f}\|_2^2 + \eta}$, with a small smoothing parameter, $\eta = 10^{-10}$. One convergence check on the algorithm is performed by evaluating

$$\cos\alpha = \frac{\left(\nabla_{\vec{f}} R(\vec{f})\right) \cdot \left(\nabla_{\vec{f}} D(X\vec{f}, \vec{g})\right)}{|\nabla_{\vec{f}} R(\vec{f})||\nabla_{\vec{f}} D(X\vec{f}, \vec{g})|} \tag{17}$$

where $R(\vec{f})$ denotes a generic regularization term, and for constrained TV-minimization $R(\vec{f}) = \|\vec{f}\|_{\mathrm{TV}}$. The conditions for convergence, derived in [4], are that the gradients of the data-error and regularization terms are back-to-back, $\cos\alpha = -1$, and $D(X\vec{f}, \vec{g}) = \epsilon$. The latter condition assumes that the data-

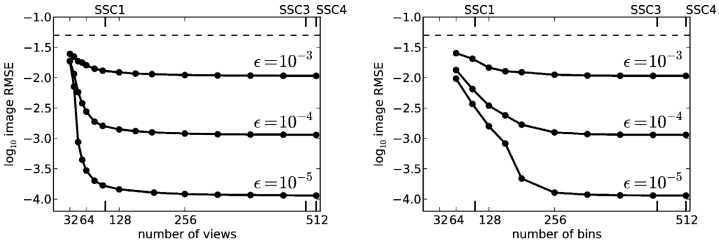


Fig. 5. Image RMSE, δ, for the data distance $D(X\vec{f},\vec{g})$ constrained by $\epsilon = 10^{-3}$, 10^{-4}, and 10^{-5}. The image size is 256 × 256. Left: fixed $N_{\text{bins}} = 2N = 512$, as function of N_{views}. Right: fixed $N_{\text{views}} = 2N = 512$, as function of N_{bins}. The labels SSC1, SSC3, and SSC4 are the sufficient-sampling conditions discussed in Section IV-A.

error constraint is active, which is the case for all the simulations performed here. For the present results, iteration is terminated when both

$$\cos\alpha \leq -0.9999$$
$$\frac{|D(X\vec{f},\vec{g}) - \epsilon|}{\epsilon} \leq 0.001 \tag{18}$$

are satisfied.

C. Admitted Undersampling by ℓ_2-TV

We are interested in two separate notions of accurate reconstruction: *exact reconstruction* and *stable reconstruction*. By the former, we mean that the reconstructed image is identical to the original. Exact reconstruction is only possible when $\epsilon = 0$, because the regularizing effect of a nonzero ϵ introduces a bias relative to the original image. Having $\epsilon = 0$ is only relevant for noise-free data, which means that stability is not an issue. Since SSC2 ensures a unique solution to (2), it can be used as a full sampling reference point for exact reconstruction. In practice, for the considered system matrix class, we found little difference in locations of SSC1 and SSC2, and we will therefore use SSC1 as a surrogate for SSC2.

Stable reconstruction is the corresponding concept for fixed, nonzero ϵ, where we cannot hope for exact reconstruction. Instead, we are interested in the degree of sampling at which further increase in sampling leads to no further improvement in the reconstruction. This point of stable reconstruction can be compared to SSC3, since that is the point where no further improvement of the system matrix condition number occurs. Since SSC4 was seen to approximate SSC3 for the present system matrix class and is simple to determine, it could also be considered to use SSC4 instead for the reference point.

In the simulations, we take the image array to be the same as that of the breast phantom $N = 256$. The parameters of the circular, fan-beam system matrix class are varied in a fashion parallel to the condition number plots of Fig. 1: first, N_{bins} is fixed at $2N$ and N_{views} is varied in the range $[32, 512]$; and second, N_{views} is fixed at $2N$ and N_{bins} is varied in the range $[32, 512]$.

Computing SSC1 and SSC4 is straightforward and they occur at $N_{\text{views}} = 101$ and $N_{\text{views}} = 512$, respectively. Computation of SSC3 was performed using the procedure outlined in Section IV-D. For the fixed-bin case, SSC3 occurs at $N_{\text{views}} = 492$, with $r_{\text{samp}} = 1.51$ and for the fixed-view case at $N_{\text{bins}} = 456$.

In order to assess the undersampling with respect to exact reconstruction admitted by exploiting image gradient-magnitude sparsity, we need access to the solution of ℓ_2-TV for a data-equality constraint, $\epsilon = 0$. We are unaware of efficient algorithms that make ℓ_2-TV with $\epsilon = 0$ feasible to solve in acceptable time for systems of realistic size. With the algorithm described in Section V-B, we can obtain an approximation by solving ℓ_2-TV with $\epsilon = 10^{-3}$, 10^{-4}, and 10^{-5} and thus study the reconstruction error as ϵ approaches zero. As an error measure, we use the root-mean-square-error (RMSE)

$$\delta = \frac{\|\vec{f}^* - \vec{f}_0\|_2}{\sqrt{N_{\text{pix}}}} \tag{19}$$

where $\vec{f}^*$ is the solution to ℓ_2-TV and $\vec{f}_0$ is the original phantom.

The computed RMSEs for the results from ℓ_2-TV are displayed in Fig. 5. As in Section IV we show SSC-locations at the top and bottom. The horizontal line shows the minimum gray level contrast, 0.05, in the test phantom and provides a reference for the RMSE. An image RMSE much less than the minimum gray-level contrast is an indicator that the reconstructed image is visually close to the original phantom (barring pathological distributions of the image error).

For the plots of δ versus N_{views}, we note a steep drop in δ as N_{views} increases past 40 views and the drop is increasingly rapid as ϵ decreases. Based on these curves we extrapolate that exact reconstruction would be attained for $N_{\text{views}} \approx 50$ at $\epsilon = 0$. Because SSC1 occurs at $N_{\text{views}} = 101$, we note an admitted undersampling with respect to exact reconstruction of a factor of 2 for the present simulation. Note that use of SSC1 leads to a conservative estimate, because SSC2 can only be larger than SSC1.

For the plots of δ versus N_{bins}, the image RMSE curves drop much more gradually at each of the ϵ's investigated. Based on these curves it is only clear that δ is tending to zero at $N_{\text{bins}} =$

190 as function of ϵ. Comparing to SSC1 at $N_{\text{bins}} = 101$, we do not observe any level of admitted undersampling in the bin-direction with respect to exact reconstruction. Extending the range of ϵ to smaller values may yield a different conclusion.

This difference reflects an asymmetry in sampling of the two parameters of X. We note that the asymmetry was also observed in the condition number dependence on N_{views} and N_{bins} for the $N = 32$ simulations in Section IV-B. For the present $N = 256$ simulations, a relatively large $\kappa(X) = 3.2 \cdot 10^4$ is seen for $N_{\text{bins}} = 128$ and $N_{\text{views}} = 512$, compared to $\kappa(X) = 1.5 \cdot 10^3$ for $N_{\text{bins}} = 512$ and $N_{\text{views}} = 128$. The results demonstrate a larger potential for successful TV-based reconstruction from few views compared to using few bins.

Regarding stable reconstruction, we note that the curves in Fig. 5 all exhibit a plateau, where δ levels off with increasing N_{views} or N_{bins}, meaning that no gain in image RMSE is achieved by increasing sampling. Thus, the left-most point of these plateaus is the point of stable reconstruction. For the plot varying N_{views}, we see that stable reconstruction begins at $N_{\text{views}} \approx 80$, which is a factor of 6 fewer than SSC3. For the plot varying N_{bins}, the stable reconstruction begins at $N_{\text{bins}} \approx 200$, a factor of approximately 2 fewer than SSC3.

These results show *quantitatively* that significant undersampling in N_{views}, particularly with respect to stable reconstruction, is admitted for ℓ_2-TV. This conclusion is achieved with a phantom modeling a realistic level of gradient-magnitude sparsity. We do point out, however, that these empirical results only apply to the presented simulation. To support the present conclusion for admitted undersampling, we vary in Section VI different aspects of the ℓ_2-TV study.

D. Altering the Optimization Problem

To support the use of the gradient-magnitude sparsity exploiting ℓ_2-TV for admitting undersampling, we compare results with two other optimization problems

$$\ell_2\text{-magnitude:} \quad \vec{f}^* = \arg\min_{\vec{f}} \|\vec{f}\|_2^2 \quad \text{subject to} \quad D(X\vec{f}, \vec{g}) \leq \epsilon \tag{20}$$

and

$$\ell_2\text{-roughness:} \quad \vec{f}^* = \arg\min_{\vec{f}} \|\nabla\vec{f}\|_2^2 \quad \text{subject to} \quad D(X\vec{f}, \vec{g}) \leq \epsilon \tag{21}$$

where ∇ represents a numerical gradient operation and is computed by forward finite-differencing. The Lagrangian form of these optimizations are two forms of Tikhonov regularization commonly used for IIR.

The solutions to ℓ_2-magnitude and ℓ_2-roughness are obtained with linear conjugate gradients (CG) applied to the Lagrangians of these problems with the multiplier being adjusted until the data-error constraint holds with equality. The convergence criteria are the same as what is specified in (18) except that $R(\vec{f}) = \|\vec{f}\|_2^2$ for ℓ_2-magnitude and $R(\vec{f}) = \|\nabla\vec{f}\|_2^2$ for ℓ_2-roughness.

We focus on the $\epsilon = 10^{-5}$ case and plot image RMSEs for ℓ_2-magnitude, ℓ_2-roughness and ℓ_2-TV as function of N_{views} for fixed $N_{\text{bins}} = 512$ in Fig. 6. We note little difference between results from ℓ_2-magnitude and ℓ_2-roughness but a large gap between these results and those of ℓ_2-TV. The optimization problems ℓ_2-TV and ℓ_2-roughness differ only on the norm of the image-gradient in the regularization term, while ℓ_2-roughness and ℓ_2-magnitude differ by the presence of the gradient. It is clear from Fig. 6 that for this simulation, the ℓ_1-norm has the greater impact.

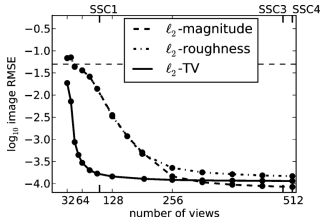


Fig. 6. Image RMSE, δ, for ℓ_2-magnitude, ℓ_2-roughness, and ℓ_2-TV reconstructions as a function of the number of views for the number of bins fixed at 512, and the data distance $D(X\vec{f}, \vec{g})$ constrained by $\epsilon = 10^{-5}$. The horizontal dashed line shows the level of the minimum gray-level contrast in the breast phantom. The labels SSC1, SSC3, and SSC4 are the sufficient-sampling conditions discussed in Section IV-A.

For large N_{views}, the ℓ_2-TV RMSE is actually slightly larger than that of ℓ_2-magnitude. The reason is the regularizing effect of having a nonzero ϵ, which causes a small bias of the solutions compared to the original image. The relative size of the biases are not known in advance. We conclude that the ℓ_2-TV solution is not to prefer over the ℓ_2-magnitude and ℓ_2-roughness solutions when N_{views} approaches the SSC3 ($r_{\text{samp}} = 1.5$). Nevertheless, there is a certain "sampling window," for the present phantom, approximately for $N_{\text{views}} \in [50, 256]$, where the TV-solution is superior to Tikhonov regularization in terms of RMSE.

In Fig. 7, we overlay the results of ℓ_2-magnitude onto the results of ℓ_2-TV from Fig. 5 in order to investigate possible undersampling admitted by ℓ_2-magnitude. The results of ℓ_2-roughness are not shown because they are similar to those of ℓ_2-magnitude and to prevent clutter in the figure. Going from left to right, both plots show a gradual decrease of δ for ℓ_2-magnitude as N_{views} and N_{bins} increase with δ leveling off at $N_{\text{views}} \approx 300$ and $N_{\text{bins}} \approx 400$. For the investigated range of ϵ, ℓ_2-magnitude does not admit any undersampling with respect to exact reconstruction, but does show a marginal undersampling with respect to stable reconstruction as the corresponding δ-curves reach the plateau before SSC3 and SSC4. In summary, the undersampling admitted by ℓ_2-magnitude is substantially less that that admitted by ℓ_2-TV for this simulation; particularly in considering view number undersampling with respect to stability.

E. Altering the Image Evaluation Metric

Conclusions based on evaluating reconstructed images with a single summarizing metric, such as the RMSE, can be misleading. While our aim is not a fully realistic image evaluation, we want to show how the results can potentially change with a

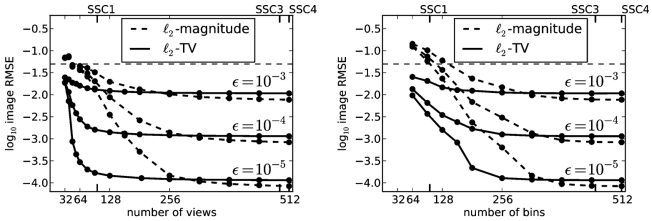


Fig. 7. Image RMSE, δ, for ℓ_2-magnitude and ℓ_2-TV reconstructions and $\epsilon = 10^{-3}, 10^{-4}$, and 10^{-5}. Left: fixed $N_{\rm bins} = 2N = 512$, as function of $N_{\rm views}$. Right: fixed $N_{\rm views} = 2N = 512$, as function of $N_{\rm bins}$. The horizontal dashed line shows the level of the minimum gray-level contrast in the breast phantom. The labels SSC1, SSC3, and SSC4 are the sufficient-sampling conditions discussed in Section IV-A.

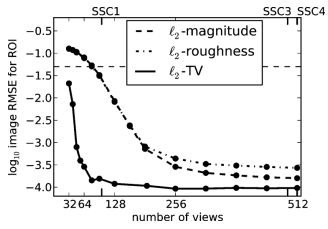


Fig. 8. Image RMSE as in Fig. 6 but for $\delta_{\rm ROI}$, i.e., the RMSE restricted to the ROI around the microcalcifications.

change of metric. For example, with the task of microcalcification detection in mind, one might consider the RMSE of only the ROI of the microcalcifications displayed in Fig. 4. This RMSE is denoted $\delta_{\rm ROI}$. Fig. 8 shows the corresponding plot for $\delta_{\rm ROI}$. While there are numerical differences between the δ and $\delta_{\rm ROI}$ plots, the trends are similar, giving us further confidence that the RMSE of the entire image, δ, can be used for investigating admitted undersampling.

Another way to evaluate images is by visual comparison. The reconstructed images in Fig. 9 are shown for a range in $N_{\rm views}$ showing the transition to accurate image reconstruction by ℓ_2-TV. That the ℓ_2-magnitude and ℓ_2-roughness show strong artifacts for this range is expected as their corresponding image RMSEs are at the level of the minimum phantom contrast level. Interestingly, the microcalcifications can be identified and well-characterized in all reconstructions, although more clearly with more views. It may be argued that, from a utility point of view, that 32 views would suffice if we are solely interested in the microcalcifications and disregard the prominent artifacts of the background image. The ROI of the ℓ_2-TV reconstructed images are shown with a narrow gray scale window in the bottom row to reveal the high level of accuracy at $N_{\rm views} = 50$. We emphasize that our goal is not a discussion about different artifacts but simply support our conclusions on undersampling from Section V-C by illustrating the behavior in the transition region around $N_{\rm views} = 50$.

F. Altering the System Matrix Class

To illustrate the change in results due to change in the system matrix class, the RMSE, δ, is again computed as function of $N_{\rm views}$ for $N_{\rm bins} = 2N = 512$ and for $\epsilon = 10^{-3}$, 10^{-4}, and 10^{-5} but using a system matrix set up through ray-tracing with nearest-neighbor interpolation at the mid-line of each pixel row, as in Section IV-C. Results are plotted in Fig. 10. The overall trends are similar to those for line-intersection, however SSC3 occurs already at $N_{\rm views} = 408$, with $r_{\rm samp} = 1.49$. By closer comparison with the line-intersection results, it is seen that the nearest-neighbor RMSEs are smaller than the line-intersection RMSE at the same $N_{\rm views}$. This example serves to illustrate that the SSCs will change when the system matrix class is altered.

While the present alteration is relatively minor, it is enough that the approximation SSC4 of to SSC3 is worse, and larger differences can be expected with more radical changes such as the use of nonpoint-like image expansion functions.

In terms of admissible undersampling for ℓ_2-TV with the altered system matrix class, we see very similar undersampling factors for both exact and stable reconstruction as for the line-intersection class.

G. Altering the Phantom Sparsity

The breast phantom study is repeated employing a variation of the FORBILD head phantom [45] which is highly sparse in the gradient-magnitude image. The present version of the phantom, which is seen in Fig. 11, does not have the ear objects of the original phantom, and the contrast levels have been increased so that the minimum gray-level contrast is the same as for the breast phantom. The gradient magnitude sparsity is 2492, or approximately a quarter of the breast phantom. In Fig. 11, the obtained δ for ℓ_2-magnitude, ℓ_2-roughness and ℓ_2-TV are shown as function of $N_{\rm views}$ with $N_{\rm bins} = 2N = 512$ and $\epsilon = 10^{-5}$.

The ℓ_2-magnitude and ℓ_2-roughness curves are almost identical to those of the breast phantom. The SSCs are unchanged, because the same system matrix class is used. That two so different looking phantoms show such similar δ-behavior suggests that the reconstruction quality of ℓ_2-magnitude and ℓ_2-roughness depend only weakly on the particular phantom.

For ℓ_2-TV, on the other hand, the step-like transition occurs already at $N_{\rm views} = 12$, for which the reconstruction is shown

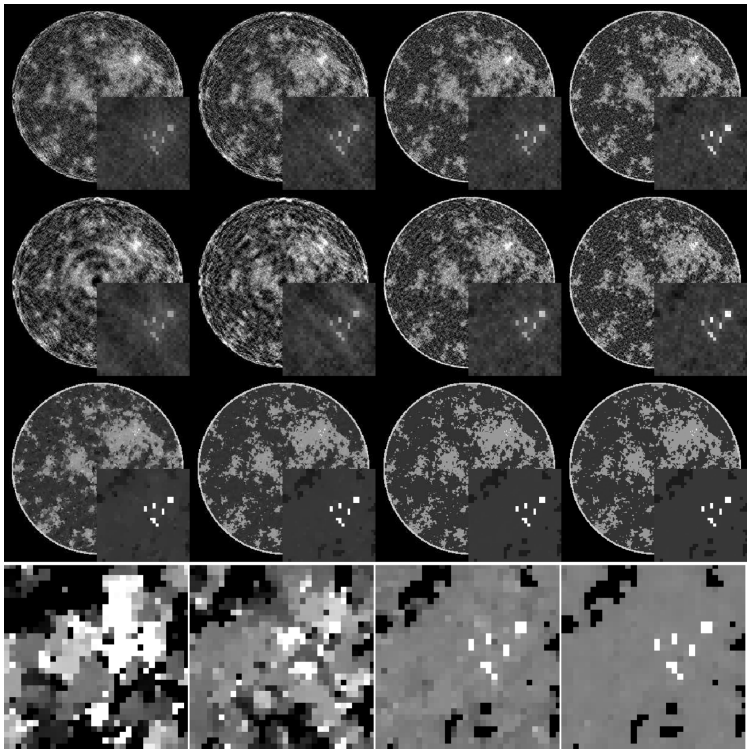


Fig. 9. First, second, third, and fourth columns show reconstructions from 32, 40, 48, 64 view data. The first, second, and third rows show ℓ_2-magnitude, ℓ_2-roughness, and ℓ_2-TV images with $\epsilon = 10^{-5}$. The gray scale window is $[0.95, 1.20]$ for the complete image, and $[0.9, 1.8]$ for the ROI blow-ups. The fourth row shows the ℓ_2-TV ROIs enlarged in an extremely narrow gray scale, $[1.09, 1.11]$, in order to scrutinize the transition to sufficient sampling based on the object sparsity. These images show that 50 views is sufficient for this system and object.

in Fig. 11. We expect that this phantom would be recovered exactly at $N_{\text{views}} = 12$ in the limit $\epsilon \to 0$, leading to an admitted undersampling with respect to exact reconstruction of a factor of approximately 8. Stable reconstruction occurs at $N_{\text{views}} \approx 64$, i.e., an undersampling also of a factor 8 with respect to stability.

Interestingly, the exact reconstruction result hints at the existence of a simple relation between sparsity and admitted undersampling. Compared to the breast phantom, there is a gain in undersampling by $8/2 = 4$. In comparison, we note that the change in gradient magnitude sparsity relative to the breast phantom is $10019/\,2492 \approx 4$. That is, reducing the sparsity by a certain factor leads to an improvement in the admitted undersampling by the same factor. This result, if shown to hold, could be important for practical application of CS-inspired sparsity-exploiting methods, since it provides quantitative insight into how many views would suffice for reconstructing images of given sparsity. Another conclusion that can be drawn is that

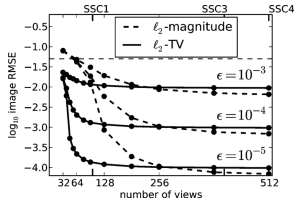


Fig. 10. Image RMSE, δ, for ℓ_2-magnitude and ℓ_2-TV reconstructions, as in Fig. 6 for varying ϵ, computing X with ray-tracing and nearest-neighbor interpolation. N_{bins} is fixed at 512. The labels SSC1, SSC3, and SSC4 are the sufficient-sampling conditions discussed in Section IV-A.

simulations with images of too low sparsity compared to a realistic level in the imaging scenario of interest are bound to yield overoptimistic promises of undersampling potential. This could

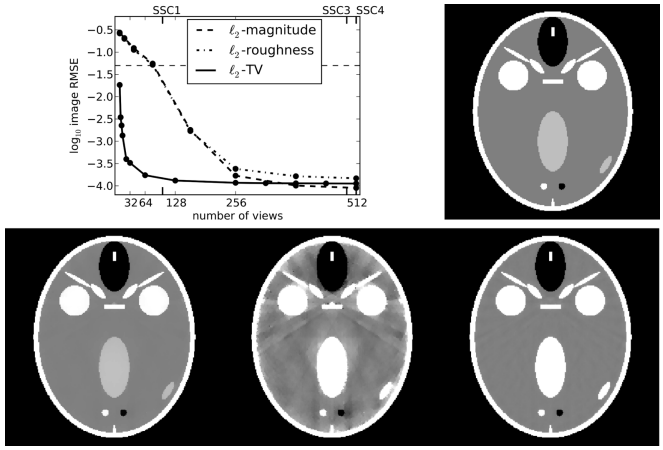


Fig. 11. Top: Image RMSE, δ, for ℓ_2-magnitude, ℓ_2-roughness, and ℓ_2-TV reconstructions, as in Fig. 6, except that the data are generated from the head phantom shown on the right. The labels SSC1, SSC3, and SSC4 are the sufficient-sampling conditions discussed in Section IV-A. Bottom: images reconstructed for $N_{\text{views}} = 12$ shown in a gray scale window of [0.9,1.1] on the left and a narrow gray scale window of $[0.99,1.01]$ in the middle. On the right is the reconstructed image for $N_{\text{views}} = 32$ in the narrow gray scale window $[0.99,1.01]$. At $N_{\text{views}} = 12$ the RMSE is 0.005 resulting in visible artifacts for the image shown in the 1% gray scale window, while the RMSE is a factor of 10 less for $N_{\text{views}} = 32$.

have been anticipated but the result establishes this expectation quantitatively.

We caution, however, that the result is based on only two phantoms and further study is required. For more robust conclusions, the present studies need to be performed on ensembles of phantoms in order to verify that admitted undersampling for constrained TV-minimization depends primarily on the gradient-magnitude image sparsity.

We also note that while we may have exact reconstruction of the head phantom at $N_{\text{views}} = 12$ and the reconstructed image at $\epsilon = 10^{-5}$ appears very accurate in the $[0.9, 1.1]$ gray scale window, it is in fact *not* an exact reconstruction. By narrowing the gray scale to $[0.99, 1.01]$, also shown in Fig. 11, prominent artifacts become visible. This underlines that exact reconstruction is not the relevant notion for a fixed, nonzero ϵ. Instead, stable reconstruction, at $N_{\text{views}} = 64$, yields an accurate reconstruction, and for the present case already at $N_{\text{views}} = 32$ (also shown in Fig. 11) the artifacts are reduced to a negligible level in the $[0.99,1.01]$ gray scale window.

VI. Summary

We argue that a quantitative notion of a sufficient-sampling condition (SSC) for X-ray CT using the DD imaging model is necessary in order to provide a reference for evaluating the undersampling potential of sparsity-exploiting methods. We propose and apply four different SSCs to a class of system matrices describing circular, fan-beam CT with a pixel expansion. While SSC1 and SSC2 characterize invertibility of the system matrix, SSC3 characterizes numerical stability for inversion of the system matrix. A simple-to-compute SSC4 is seen to approximate SSC3 closely for the circular, fan-beam full angular range CT geometry.

We employ the SSCs as reference points of full sampling to quantify undersampling admitted by reconstruction through TV-minimization on a breast CT simulation. Relative to SSC1, we observe some undersampling potential of TV-minimization for exact reconstruction and large undersampling relative to SSC3 and SSC4 for stable reconstruction. We find few-view reconstruction to admit larger undersampling than few-detector bin reconstruction, and we show evidence of a simple quantitative relation between image sparsity and the admitted undersampling.

More generally, the proposed SSCs can help to engineer and understand other sparsity-exploiting IIR algorithms by providing full sampling reference points for quantification of admissible undersampling in other imaging applications of interest. This analysis can guide decisions on alternative optimization problems, object representations, sampling configurations, and integration models.

Acknowledgment

The authors are very grateful to the referees for providing valuable feedback that significantly improved the quality of the paper.

References

[1] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-ray Sci. Technol.*, vol. 14, pp. 119–139, 2006.

[2] J. Song, Q. H. Liu, G. A. Johnson, and C. T. Badea, "Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT," *Med. Phys.*, vol. 34, pp. 4476–4483, 2007.

[3] G.-H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, pp. 660–663, 2008.

[4] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, pp. 4777–4807, 2008.

[5] E. Y. Sidky, X. Pan, I. S. Reiser, R. M. Nishikawa, R. H. Moore, and D. B. Kopans, "Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms," *Med. Phys.*, vol. 36, pp. 4920–4932, 2009.

[6] X. Pan, E. Y. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?," *Inverse Prob.*, vol. 25, pp. 123009–123009, 2009.

[7] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT," *Phys. Med. Biol.*, vol. 55, pp. 6575–6599, 2010.

[8] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrieß, "Improved total variation-based CT image reconstruction applied to clinical data," *Phys. Med. Biol.*, vol. 56, pp. 1545–1561, 2011.

[9] X. Han, J. Bian, D. R. Eaker, T. L. Kline, E. Y. Sidky, E. L. Ritman, and X. Pan, "Algorithm-enabled low-dose micro-CT imaging," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 606–620, Mar. 2011.

[10] M. Defrise, C. Vanhove, and X. Liu, "An algorithm for total variation regularization in high-dimensional linear problems," *Inverse Prob.*, vol. 27, p. 065002, 2011.

[11] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, 2006.

[12] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[13] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 21, no. 2, pp. 21–30, Mar. 2008.

[14] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, pp. 1207–1223, 2006.

[15] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

[16] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: An application to limited-angle tomography," *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 204–221, Feb. 1998.

[17] M. Persson, D. Bone, and H. Elmqvist, "Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography," *Phys. Med. Biol.*, vol. 46, pp. 853–866, 2001.

[18] I. A. Elbakri and J. A. Fessler, "Statistical image reconstruction for polyenergetic X-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, Feb. 2002.

[19] M. Li, H. Yang, and H. Kudo, "An accurate iterative reconstruction algorithm for sparse objects: Application to 3-D blood vessel reconstruction from a limited number of projections," *Phys. Med. Biol.*, vol. 47, pp. 2599–2609, 2002.

[20] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.

[21] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, "Implementation of an optimal first-order method for strongly convex total variation regularization," *BIT*, vol. 52, pp. 329–356, 2012.

[22] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, pp. 120–145, 2011.

[23] S. Ramani and J. A. Fessler, "A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 677–688, Mar. 2012.

[24] E. A. Rashed and H. Kudo, "Statistical image reconstruction from limited projection data with intensity priors," *Phys. Med. Biol.*, vol. 57, pp. 2039–2061, 2012.

[25] E. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Phys. Med. Biol.*, vol. 57, pp. 3065–3091, 2012.

[26] H. H. Barrett and K. J. Myers, *Foundations of Image Science*. Hoboken, NJ: Wiley, 2004.

[27] S. Matej and R. M. Lewitt, "Practical considerations for 3-D image reconstruction using spherically symmetric volume elements," *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 68–78, Feb. 1996.

[28] A. H. Delaney and Y. Bresler, "Multiresolution tomographic reconstruction using wavelets," *IEEE Trans. Image Process.*, vol. 4, no. 6, pp. 799–813, Jun. 1995.

[29] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc. IEEE*, vol. 84, pp. 626–638, 1996.

[30] M. H. Buonocore, W. R. Brody, and A. Macovski, "A natural pixel decomposition for two-dimensional image reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 28, no. 2, pp. 69–78, Feb. 1981.

[31] G. T. Gullberg and G. L. Zeng, "A reconstruction algorithm using singular value decomposition of a discrete representation of the exponential Radon transform using natural pixels," *IEEE Trans. Nucl. Sci.*, vol. 41, no. 6, pp. 2812–2819, Dec. 1994.

[32] F. Natterer, *The Mathematics of Computerized Tomography*. New York, NY: Wiley, 1986.

[33] F. Natterer, "Sampling in fan beam tomography," *SIAM J. Appl. Math.*, vol. 53, pp. 358–380, 1993.

[34] S. H. Izen, D. P. Rohler, and K. L. A. Sastry, "Exploiting symmetry in fan beam CT: Overcoming third generation undersampling," *SIAM J. Appl. Math.*, vol. 65, pp. 1027–1052, 2005.

[35] A. Faridani, "Fan-beam tomography and sampling theory," in *Proc. Symp. Appl. Math.*, Atlanta, GA, 2006, pp. 43–66, Amer. Mathematical Soc..

[36] H. H. Barrett and H. Gifford, "Cone-beam tomography with discrete data sets," *Phys. Med. Biol.*, vol. 39, pp. 451–476, 1994.

[37] H. H. Barrett, J. L. Denny, R. F. Wagner, and K. J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," *J. Opt. Soc. Amer. A*, vol. 12, pp. 834–852, 1995.

[38] P. J. L. Rivière and X. Pan, "Sampling and aliasing consequences of quarter-detector offset use in helical CT," *IEEE Trans. Med. Imag.*, vol. 23, no. 6, pp. 738–749, Jun. 2004.

[39] R. H. Huesman, "The effects of a finite number of projection angles and finite lateral sampling of projections on the propagation of statistical errors in transverse section reconstruction," *Phys. Med. Biol.*, vol. 22, pp. 511–521, 1977.

[40] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*. Manchester, U.K.: Manchester Univ. Press, 1992.

[41] I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," *Med. Phys.*, vol. 37, pp. 1591–1600, 2010.

[42] J. M. Boone, T. R. Nelson, K. K. Lindfors, and J. A. Seibert, "Dedicated breast CT: Radiation dose and image quality evaluation," *Radiology*, vol. 221, pp. 657–667, 2001.

[43] S. J. Glick, "Breast CT," *Annu. Rev. Biomed. Eng.*, vol. 9, pp. 501–526, 2007.

[44] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*. New York: Springer, 2005.

[45] G. Lauritsch and H. Bruder, Forbild head phantom [Online]. Available: http://www.imp.uni-erlangen.de/phantoms/head/head.html

Appendix B

First-order convex feasibility algorithms for x-ray CT

Med. Phys., vol. 40, issue 3, p. 031115, 2013.
doi:10.1118/1.4790698.
Published 28 February 2013.

E. Y. Sidky, **J. S. Jørgensen** and X. Pan

First-order convex feasibility algorithms for x-ray CT

Emil Y. Sidky[a)]
Department of Radiology, University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois, 60637

Jakob S. Jørgensen[b)]
Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet, Building 303B, 2800 Kongens Lyngby, Denmark

Xiaochuan Pan[c)]
Department of Radiology, University of Chicago, 5841 South Maryland Avenue, Chicago, Illinois, 60637

(Received 13 August 2012; revised 30 November 2012; accepted for publication 23 January 2013; published 28 February 2013)

Purpose: Iterative image reconstruction (IIR) algorithms in computed tomography (CT) are based on algorithms for solving a particular optimization problem. Design of the IIR algorithm, therefore, is aided by knowledge of the solution to the optimization problem on which it is based. Often times, however, it is impractical to achieve accurate solution to the optimization of interest, which complicates design of IIR algorithms. This issue is particularly acute for CT with a limited angular-range scan, which leads to poorly conditioned system matrices and difficult to solve optimization problems. In this paper, we develop IIR algorithms which solve a certain type of optimization called convex feasibility. The convex feasibility approach can provide alternatives to unconstrained optimization approaches and at the same time allow for rapidly convergent algorithms for their solution—thereby facilitating the IIR algorithm design process.
Methods: An accelerated version of the Chambolle−Pock (CP) algorithm is adapted to various convex feasibility problems of potential interest to IIR in CT. One of the proposed problems is seen to be equivalent to least-squares minimization, and two other problems provide alternatives to penalized, least-squares minimization.
Results: The accelerated CP algorithms are demonstrated on a simulation of circular fan-beam CT with a limited scanning arc of 144°. The CP algorithms are seen in the empirical results to converge to the solution of their respective convex feasibility problems.
Conclusions: Formulation of convex feasibility problems can provide a useful alternative to unconstrained optimization when designing IIR algorithms for CT. The approach is amenable to recent methods for accelerating first-order algorithms which may be particularly useful for CT with limited angular-range scanning. The present paper demonstrates the methodology, and future work will illustrate its utility in actual CT application. [http://dx.doi.org/10.1118/1.4790698]

Key words: computed tomography, iterative image reconstruction, convex optimization

I. INTRODUCTION

Iterative image reconstruction (IIR) algorithms in computed tomography (CT) are designed based on some form of optimization. When designing IIR algorithms to account for various factors in the CT model, the actual designing occurs usually at the optimization problem and not the individual processing steps of the IIR algorithm. Once the optimization problem is established, algorithms are developed to solve it. Achieving convergent algorithms is important, because they yield access to the designed solution of the optimization problem and allow for direct assessment of what factors to include in a particular optimization. Convergent algorithms can also aid in determining at what iteration number to truncate an IIR algorithm. With access to the designed solution, the difference between it and previous iterates can be quantitatively evaluated to see whether this difference is significant with respect to a given CT imaging task.

It can be challenging to develop convergent algorithms for some optimization problems of interest. This issue is particularly acute for CT, which involves large-scale optimization. In using the term "large-scale," we are specifically referring to optimization problems based on a linear data model, and the dimension of the linear system is so large that the system matrix cannot be explicitly computed and stored in memory. Such systems only allow for algorithms which employ operations of a similar computational expense to matrix-vector products. Large-scale optimization algorithms are generally restricted to first-order methods, where only gradient information on the objective is used, or row-action algorithms such as the algebraic reconstruction technique (ART).[1,2] Recently, there has been renewed interest in developing convergent algorithms for optimization problems involving ℓ_1-based image norms, and only in the last couple of years have practical, convergent algorithms been developed to solve these optimization problems for IIR in CT.[3–6]

Despite the progress in algorithms, there are still CT configurations of practical interest, which can lead to optimization problems that can be quite challenging to solve accurately. Of particular interest in this work is CT with a limited angular-range scanning arc. Such a configuration is relevant to many C-arm CT and tomosynthesis applications. Modeling limited angular-range scanning leads to system matrices with unfavorable singular value spectra and optimization problems for which many algorithms converge slowly.

In this paper, we consider application of convex feasibility[7,8] to IIR for CT. In convex feasibility, various constraints on properties of the image are formulated so that each of these constraints specifies a convex set. Taking the intersection of all of the convex sets yields a single convex set, and the idea is to simply choose one of these images in the intersection set. We have found convex feasibility to be useful for CT IIR algorithm design,[9] and it is of particular interest here for limited angular-range CT, because convex feasibility is amenable to recent accelerated first-order algorithms proposed by Chambolle and Pock (CP).[10] In Sec. II, we specify the limited angular-range CT system, discuss unconstrained optimization approaches, and then list three useful convex feasibility problems along with a corresponding accelerated CP algorithm. In Sec. III, the accelerated convex feasibility CP algorithms are demonstrated with simulated CT projection data.

II. METHODS: CHAMBOLLE–POCK ALGORITHMS FOR CONVEX FEASIBILITY

For this paper, we focus on modeling circular, fan-beam CT with a limited scanning angular range. As with most work on IIR, the data model is discrete-to-discrete (DD) and can be written as a linear equation

$$\mathbf{g} = X\mathbf{f}, \tag{1}$$

where $\mathbf{f}$ is the image vector composed of pixel coefficients, X is the system matrix generated by computing the ray-integrals with the line-intersection method, and $\mathbf{g}$ is the data vector containing the estimated projection samples. For the present investigation on IIR algorithms, we consider a single configuration for limited angular range scanning where the system matrix X has a left-inverse ($X^T X$ is invertible) but is numerically unstable in the sense that it has a large condition number. The vector $\mathbf{f}$ consists of the pixels within a circle inscribed in a 256×256 pixel array; the total number of pixels is 51 468. The sinogram contains 128 views spanning a 144° scanning arc, and the projections are taken on a 512-bin linear detector array. The modeled source-to-isocenter and source-to-detector distances are 40 and 80 cm, respectively. The total number of transmission measurements is 65 536, and as a result the system matrix X has about 25% more rows than columns. The condition of X, however, is poor, which can be understood by considering the corresponding continuous-to-continuous (CC) fan-beam transform. A sufficient angular range for stable inversion of the CC fan-beam transform requires a 208° scanning arc (180° plus the fan-angle, see, for example, Sec. 3.5 of Ref. 11). By using the inverse power method, as described in Ref. 12, the condition number, the ratio of the largest to smallest singular value, for X is determined to be 2.55×10^4. One effect of the large condition number is to amplify noise present in the data, but it can also cause slow convergence for optimization-based IIR.

II.A. Unconstrained optimization for IIR in CT

Image reconstruction using this DD data model is usually performed with some form of optimization, because physical factors and inaccuracy of the model render Eq. (1) inconsistent, namely, no $\mathbf{f}$ exists satisfying this equation. Typically in using this model, quadratic optimization problems are formulated, the simplest of which is the least-squares problem

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{g} - X\mathbf{f}\|_2^2 \right\}, \tag{2}$$

where $\mathbf{f}^\circ$ is the image which minimizes the Euclidean distance between the available data $\mathbf{g}$ and the estimated data $X\mathbf{f}$. In the remainder of the paper, we use the superscript "°" to indicate a solution to an optimization problem. Taking the gradient of this objective, and setting it to zero componentwise, leads to the following consistent linear equation

$$X^T X\mathbf{f} = X^T \mathbf{g}, \tag{3}$$

where the superscript T denotes the matrix transpose. This linear equation is particularly useful for setting up the linear conjugate gradients (CG) algorithm, see for example Ref. 13, which has been used as the gold standard algorithm for large-scale quadratic optimization in IIR. The reader is also referred to conjugate gradients least-squares (CGLS) and LSQR (an algorithm for sparse linear equations and sparse least squares), which solve Eq. (2) for nonsymmetric X.[14]

The solution to Eq. (2) or (3) can be undesirable because of inconsistency in the data. Particularly for the present case, the poor conditioning of X can yield tremendously amplified artifacts in the reconstructed image. As is well-known, artifacts due to data inconsistency can be controlled in optimization-based IIR by adding a penalty term to discourage large variations between neighboring pixels

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{g} - X\mathbf{f}\|_2^2 + \alpha R(\mathbf{f}) \right\}, \tag{4}$$

where $R(\mathbf{f})$ is a generic roughness term which usually is a convex function of the difference between neighboring pixels in the image. The parameter α controls the strength of the penalty with larger values leading to smoother images. When $R(\mathbf{f})$ is chosen to be quadratic in the pixel values, the optimization problem can be solved by a host of standard algorithms including CG. Of recent interest have been convex regularizers based on the ℓ_1-norm, which is more difficult to treat and, accordingly, for which many new, convergent algorithms have been proposed and applied to image reconstruction in CT.[3–6]

II.B. Convex feasibility

In this paper, we consider convex feasibility problems which provide alternatives to the above-mentioned

optimization problems. For convex feasibility problems, convex sets resulting from constraints on various properties of the image are formulated, and a single image which satisfies all the imposed constraints is sought. Most algorithms for such problems are based on projection onto convex sets (POCS),[8] where the image estimate is sequentially projected onto each constraint set. Convex feasibility problems can be: inconsistent, no image satisfies all the constraints; or consistent, at least one image satisfies all the constraints. In either case, POCS algorithms can yield a useful solution. In the inconsistent case, POCS algorithms can be designed to yield an image "close" to satisfying all the constraints. In the consistent case, a POCS algorithm can be designed to find an image obeying all the constraints. In either case, the issue of uniqueness is secondary, as an image "in the middle" of many inconsistent constraints or in the intersection set of consistent constraints is considered to be equally valid. Accordingly, the POCS result often depends on starting image, relaxation schemes, and projection order.

For our purposes, we write a general convex feasibility as the following optimization

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \sum_i \delta_{S_i}(K_i(\mathbf{f})) \right\}, \tag{5}$$

$K_i(\cdot)$ is the ith affine transform of the image $\mathbf{f}$; S_i is the ith convex set to which $K_i(\mathbf{f})$ belongs; and the indicator function δ is defined

$$\delta_S(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in S \\ \infty & \mathbf{x} \notin S \end{cases}. \tag{6}$$

The use of indicator functions in convex analysis provides a means to turn convex sets into convex functions,[15] and in this case, they allow convex feasibility problems to be written as a minimization of a single objective function. The objective function in Eq. (5) is zero for any image $\mathbf{f}$ satisfying all the constraints, i.e., $K_i(\mathbf{f}) \in S_i$ for all i, and it is infinity if any of the constraints are violated. For a consistent convex feasibility problem, the objective minimum is zero, and for an inconsistent convex feasibility problem, the objective minimum is infinity.

II.C. Modified convex feasibility optimization and the Chambolle–Pock primal–dual algorithm

To solve the generic convex feasibility problem in Eq. (5), we modify this optimization by adding a quadratic term

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \sum_i \delta_{S_i}(K_i(\mathbf{f})) \right\}, \tag{7}$$

where $\mathbf{f}_{\text{prior}}$ is a prior image estimate that can be set to zero if no prior image is available. With this optimization problem, we actually specify a unique solution to our generic convex feasibility problem in the consistent case, namely, the image satisfying all constraints and closest to $\mathbf{f}_{\text{prior}}$. As we will demonstrate the algorithm we propose to use for solving Eq. (7) appears to yield useful solutions for the *inconsistent* case. This latter property can be important for IIR in CT because the data model in Eq. (1) is often inconsistent with the available projection data.

The reason for recasting the optimization in the form shown in Eq. (7) is that this optimization can be solved by an accelerated algorithm described in Ref. 10. Recently, we have been interested in a convex optimization framework and algorithms derived by Chambolle and Pock (CP).[10,16] This framework centers on the generic convex optimization

$$p^\circ = \min_{\mathbf{x}} \{G(\mathbf{x}) + F(H\mathbf{x})\}, \tag{8}$$

where $G(\cdot)$ and $F(\cdot)$ are convex functions, and H is a linear transform. The objective function

$$p = G(\mathbf{x}) + F(H\mathbf{x})$$

is referred to as the primal objective. This generic problem encompasses many optimizations of interest to IIR in CT, because nonsmooth convex functions such as the indicator and ℓ_1-norm can be incorporated into F or G. Also, the linear transform H can model projection, for a data fidelity term, or a finite-difference-based gradient, for an image total variation (TV) term. The CP framework, as presented in Ref. 10, comes with four algorithms that have different worst-case convergence rates depending on convexity properties of F and G. Let N be the number of iterations, the algorithm summaries are:

- **CP Algorithm 1:** This basic CP algorithm forms the basis of the subsequent algorithms and it only requires F and G to be convex. The worst-case convergence rate is $O(1/N)$.
- **CP Algorithm 2:** Can be used if either F or G are uniformly convex. Modifies CP Algorithm 1 using a step-size formula developed by Nesterov.[17,18] The worst-case convergence rate is $O(1/N^2)$. Because the convergence rate is faster than the previous case, this algorithm is an accelerated version of CP Algorithm 1.
- **CP Algorithm 3:** Can be used if both F and G are uniformly convex. This algorithm is the same as CP Algorithm 1, except that there is a specific choice of algorithm parameters, depending on constants related to the uniform convexity of F and G. The worst-case convergence is linear, i.e., $O(1/c^N)$, where $c > 1$ is a constant.
- **CP Algorithm 4:** A simpler version of CP Algorithm 2, which also requires F or G to be uniformly convex. The convergence rate is slightly worse than $O(1/N^2)$.

In a previous publication,[6] we illustrated how to use CP Algorithm 1 from Ref. 10 to prototype many optimization problems of potential interest to image reconstruction in CT. We were restricted to CP Algorithm 1, because we considered mainly problem where G was 0, and F contained indicator functions, the ℓ_1-norm, or TV terms and accordingly F was not uniformly convex. In the present work, we narrow the class of optimization problems to those which can be written in the form of Eq. (7), where the sets S_i are simple enough that direct Euclidean projections to the sets S_i are analytically available. In matching up Eq. (7) to the generic optimization in Eq. (8), the function G is assigned the uniformly convex quadratic term and F gets the sum of indicator functions. As

such, Eq. (7) fills the requirements of CP Algorithms 2 and 4. In the particular case of Eq. (7) the uniformly convex term, $0.5\|\mathbf{f}-\mathbf{f}_{\text{prior}}\|_2^2$, is simple enough that CP Algorithm 2 can be derived without any difficulty. Because this algorithm is an accelerated version of CP Algorithm 1, we refer to it, here, as the accelerated CP algorithm. This algorithm acceleration is particularly important for IIR involving an ill-conditioned data model such as Eq. (1) in the case of limited angular range scanning.

II.D. The primal-dual gap and convergence criteria

The CP algorithms are primal-dual in that they solve the primal minimization Eq. (8) together with a dual maximization

$$d^\circ = \max_{\mathbf{y}} \{-F^*(\mathbf{y}) - G^*(-H^T\mathbf{y})\}, \tag{9}$$

and,

$$d = -F^*(\mathbf{y}) - G^*(-H^T\mathbf{y})$$

is the dual objective, and the superscript * represents convex conjugation through the Legendre transform

$$P^*(\mathbf{z}) = \max_{\mathbf{z}'} \{\mathbf{z}^T\mathbf{z}' - P(\mathbf{z}')\}. \tag{10}$$

That the CP algorithms obtain the dual solution, also, is useful for obtaining a robust convergence criterion that applies for nonsmooth convex optimization. As long as the primal objective p is convex, we have $p^\circ = d^\circ$. While a solution for a smooth optimization can be checked by observing that the gradient of the primal objective in Eq. (8) is zero, this test may not be applicable to nonsmooth optimization, where the primal objective may not be differentiable at its minimum. Instead, we can use the primal-dual gap $p - d$, because the primal objective for any $\mathbf{x}$ is larger than the dual objective in Eq. (9) for any $\mathbf{y}$ except when $\mathbf{x}$ and $\mathbf{y}$ are at their respective extrema, where these objectives are equal. Checking the primal-dual gap is complicated slightly when indicator functions are included in one of the objectives, because indicators take on infinite values when their corresponding constraint is not satisfied. As a result, we have found it convenient in Ref. 6 to define a conditional primal-dual gap which is the primal-dual gap with indicator functions removed from both objectives. This convergence check then involves observing that the conditional primal-dual gap is tending to zero and that the iterates are tending toward satisfying each of the constraints corresponding to the indicator functions. By dividing up the convergence check in this way, we give up nonnegativity of the gap. The conditional primal-dual gap can be negative, but it will approach zero as the iterates approach the solution to their respective optimizations. Use of this convergence check will become more clear in the results section where it is applied to various convex feasibility problems related to IIR in CT.

With respect to numerical convergence, it is certainly useful to have mathematical convergence criteria such as the gradient of the objective or the primal-dual gap, but it is also important to consider metrics of interest. By a metric, we mean some function of the image pixel values pertaining to a particular purpose or imaging task. For numerical convergence, we need to check, both, that the convergence metrics are approaching zero and that other metrics of interest are leveling off so that they do not change with further iterations. Rarely are IIR algorithms run to the point where the convergence criterion are met exactly, in the numerical sense. This means, that the image estimates are still evolving up until the last computed iteration, and one cannot say *a priori* whether the small changes in the image estimates are important to the metrics of interest or not. For the present theoretical work, where we have access to the true underlying image, we employ the image root mean square error (RMSE) as an image quality metric. But we point out that other metrics may be more sensitive and potentially alter the iteration number where the specific problem can be considered as converged.[19]

II.E. Convex feasibility instances

In the following, we write various imaging problems in the form of Eq. (7). We consider the following three convex feasibility problems: EC, one set specifying a data equality constraint; IC, one set specifying a data inequality constraint; and ICTV, two sets specifying data and TV inequality constraints. The derived accelerated CP algorithms for each problem are labeled CP2-EC, CP2-IC, and CP2-ICTV, respectively. Using simulated fan-beam CT data with a limited angular-range scanning arc, Sec. III presents results for all three problems in the consistent case and problems EC and ICTV in the inconsistent case. Of particular importance, CP2-EC applied to the inconsistent case appears to solve the ubiquitous least-squares optimization with a convergence rate competitive with CG.

II.E.1. CP2-EC: An accelerated CP algorithm instance for a data equality constraint

The data model in Eq. (1) cannot be used directly as an implicit imaging model for real CT data, because inconsistencies inherent in the data prevent a solution. But treating this equation as an implicit imaging model for ideal simulation can be useful for algorithm comparison and testing implementations of the system matrix X; we use it for the former purpose. We write this ideal imaging problem into an instance of Eq. (7),

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2}\|\mathbf{f}-\mathbf{f}_{\text{prior}}\|_2^2 + \delta_0(X\mathbf{f}-\mathbf{g}) \right\}, \tag{11}$$

where the indicator $\delta_0(\cdot)$ is zero only when all components of the argument vector are zero, and otherwise it is infinity. The corresponding dual maximization needed for computing the conditional primal-dual gap is

$$\mathbf{y}^\circ = \arg\max_{\mathbf{y}} \left\{ -\frac{1}{2}\|X^T\mathbf{y}\|_2^2 - \mathbf{g}^T\mathbf{y} + \mathbf{f}_{\text{prior}}^T(X^T\mathbf{y}) \right\}. \tag{12}$$

In matching Eq. (11) with Eq. (7), there is only one convex constraint where $K_1(\mathbf{f}) = X\mathbf{f}-\mathbf{g}$ and S_1 is the 0-vector with size, size($\mathbf{g}$). In considering ideal data and a left-invertible system matrix X, there is only one image for which the indicator is not infinite. In this situation, the first quadratic has

1: $L \leftarrow \|X\|_2$; $\tau \leftarrow 1$; $\sigma \leftarrow 1/L^2$; $n \leftarrow 0$
2: initialize $\mathbf{f}_0$ and $\mathbf{y}_0$ to zero vectors
3: $\bar{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$
4: **repeat**
5: $\quad \mathbf{y}_{n+1} \leftarrow \mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g})$
6: $\quad \mathbf{f}_{n+1} \leftarrow \left[\mathbf{f}_n - \tau(X^T\mathbf{y}_{n+1} - \mathbf{f}_{\text{prior}})\right]/(1+\tau)$
7: $\quad \theta \leftarrow 1/\sqrt{1+2\tau}$; $\tau \leftarrow \tau\theta$; $\sigma \leftarrow \sigma/\theta$
8: $\quad \bar{\mathbf{f}}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$
9: $\quad n \leftarrow n+1$
10: **until** $n \geq N$

FIG. 1. Pseudocode for N steps of the accelerated CP algorithm instance for solving Eq. (11). Variables are defined in the text.

1: $L \leftarrow \|X\|_2$; $\tau \leftarrow 1$; $\sigma \leftarrow 1/L^2$; $n \leftarrow 0$
2: initialize $\mathbf{f}_0$ and $\mathbf{y}_0$ to zero vectors
3: $\bar{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$
4: **repeat**
5: $\quad \mathbf{y}'_n \leftarrow \mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g})$; $\mathbf{y}_{n+1} \leftarrow \max(\|\mathbf{y}'_n\|_2 - \sigma\epsilon', 0)\frac{\mathbf{y}'_n}{\|\mathbf{y}'_n\|_2}$
6: $\quad \mathbf{f}_{n+1} \leftarrow \left[\mathbf{f}_n - \tau(X^T\mathbf{y}_{n+1} - \mathbf{f}_{\text{prior}})\right]/(1+\tau)$
7: $\quad \theta \leftarrow 1/\sqrt{1+2\tau}$; $\tau \leftarrow \tau\theta$; $\sigma \leftarrow \sigma/\theta$
8: $\quad \bar{\mathbf{f}}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$
9: $\quad n \leftarrow n+1$
10: **until** $n \geq N$

FIG. 2. Pseudocode for N steps of the accelerated CP algorithm instance for solving Eq. (13) with parameter ϵ'. Variables are defined in Sec. II.E.1.

no effect on the solution and accordingly the solution is independent of the prior image estimate $\mathbf{f}_{\text{prior}}$. If the system matrix is not left-invertible, the solution to Eq. (11) is the image satisfying Eq. (1) closest to $\mathbf{f}_{\text{prior}}$.

Following the formalism of Ref. 10, we write an accelerated CP algorithm instance for solving Eq. (11) and its dual Eq. (12) in Fig. 1. We define the pseudocode variables and operations starting from the first line. The variable L is assigned the matrix ℓ_2-norm of X, which is its largest singular value. This quantity can be computed by the standard power method, see Ref. 6 for its application in the present context. The parameters τ and σ control the step sizes in the primal and dual problems, respectively, and they are initialized so that their product yields $1/L^2$. Other choices on how to balance the starting values of τ and σ can be made, but we have found that the convergence of our examples does not depend strongly on the choice of these parameters. Line 5 shows the update of the dual variable $\mathbf{y}_{n+1}$; this variable has the same dimension as the data vector $\mathbf{g}$. Line 6 updates the image, and Line 7 adjusts the step-sizes in a way that accelerates the CP algorithm.[10]

II.E.2. CP2-IC: An accelerated CP algorithm instance for inequality constrained data-error

Performing IIR with projection data containing inconsistency requires some form of image regularization. One common strategy is to employ Tikhonov regularization, see for example, Chap. 2 of Ref. 20. Tikhonov regularization fits into the form of Eq. (4) by writing $R(\mathbf{f}) = (1/2)\|\mathbf{f}\|_2^2$. One small inconvenience with this approach, however, is that the physical units of the two terms in the objective of Eq. (4) are different, and therefore it can be difficult to physically interpret the regularization parameter α. An equivalent optimization can be formulated as a special case of Eq. (7),

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \delta_{\text{Ball}(\epsilon')}(X\mathbf{f} - \mathbf{g}) \right\}, \tag{13}$$

which differs from Eq. (11) only in that the set S_1 is widened from a 0-vector to Ball(ϵ'), where we use the term Ball(ϵ') to denote a multidimensional solid sphere of radius ϵ' and the dimension of the solid sphere is taken to be the same as size($\mathbf{g}$). We also define the parameter ϵ, which is a constraint on the data RMSE,

$$\epsilon = \epsilon'/\sqrt{\text{size}(\mathbf{g})}.$$

The corresponding dual maximization is

$$\mathbf{y}^\circ = \arg\max_{\mathbf{y}} \left\{ -\frac{1}{2}\|X^T\mathbf{y}\|_2^2 - \epsilon'\|\mathbf{y}\|_2 - \mathbf{g}^T\mathbf{y} + \mathbf{f}_{\text{prior}}^T(X^T\mathbf{y}) \right\}. \tag{14}$$

The indicator $\delta_{\text{Ball}(\epsilon')}(X\mathbf{f} - \mathbf{g})$ in Eq. (13) is zero when $\|X\mathbf{f} - \mathbf{g}\|_2 \leq \epsilon'$ and infinity otherwise. This optimization is equivalent to Tikhonov regularization when $\mathbf{f}_{\text{prior}}$ is zero and $\epsilon' > 0$ in the sense that there exists a corresponding α (not known ahead of time) where the two optimizations yield the same solution. The advantage of Eq. (13) is that the parameter ϵ' has a meaningful physical interpretation as a tolerance on the data-error. Larger ϵ' yields greater regularization. Generally, the Tikhonov form is preferred due to algorithm availability. Tikhonov regularization can be solved, for example, by linear CG. With the application of CP2-IC, however, an accelerated solver is now available that directly solves the constrained minimization in Eq. (13).

The pseudocode for CP2-IC is given in Fig. 2. This pseudocode differs from the previous at the update of the dual variable $\mathbf{y}_{n+1}$ in Line 5. The derivation of this dual update is covered in detail in our previous work on the application of the CP algorithm to CT image reconstruction.[6] For the limited angular-range CT problem considered here, Eq. (13) is particularly challenging because the constraint shape is highly eccentric due to the spread in singular values of X.

II.E.3. CP2-ICTV: An accelerated CP algorithm instance for total variation and data-error constraints

Recently, regularization based on the ℓ_1-norm has received much attention. In particular, the TV seminorm has found extensive application in medical imaging due to the fact that tomographic images are approximately piecewise constant. The TV seminorm of $\mathbf{f}$ is written as $\|(|\nabla\mathbf{f}|)\|_1$, where ∇ is a

matrix encoding a finite-difference approximation to the gradient operator; it acts on an image and yields a spatial-vector image. The absolute value operation acts pixelwise, taking the length of the spatial-vector at each pixel of this image; accordingly, $|\nabla \mathbf{f}|$ is the gradient-magnitude image of $\mathbf{f}$. The TV seminorm can be used as a penalty with the generic optimization of Eq. (4), by setting $R(\mathbf{f}) = \|(|\nabla \mathbf{f}|)\|_1$. Convergent large-scale solvers for this optimization problem have only recently been developed with some algorithms relying on smoothing the TV term.[3–5] As with Tikhonov regularization, there is still the inconvenience of having no physical meaning of the regularization parameter α. We continue along the path of recasting optimization problems as a convex feasibility problem and consider

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \delta_{\text{Ball}(\epsilon')}(X\mathbf{f} - \mathbf{g}) + \delta_{\text{Diamond}(\gamma)}(|\nabla \mathbf{f}|) \right\}, \tag{15}$$

where the additional indicator places a constraint on the TV of $\mathbf{f}$; and we have $K_1(\mathbf{f}) = X\mathbf{f} + \mathbf{g}$, $K_2(\mathbf{f}) = \nabla \mathbf{f}$, $S_1 = \{\mathbf{g} \text{ such that } \mathbf{g} \in \text{Ball}(\epsilon')\}$, and $S_2 = \{\mathbf{z} \text{ such that } |\mathbf{z}| \in \text{Diamond}(\gamma)\}$, where $\mathbf{z}$ is a spatial-vector image. The term Diamond(γ) describes the ℓ_1-ball of scale γ; the indicator $\delta_{\text{Diamond}(\gamma)}(|\nabla \mathbf{f}|)$ is zero when $\|(|\nabla \mathbf{f}|)\|_1 \leq \gamma$. This convex feasibility problem asks for the image that is closest to $\mathbf{f}_{\text{prior}}$ and satisfies the ϵ'-data-error and γ-TV-constraints. The corresponding dual maximization is

$$\mathbf{y}^\circ = \arg\max_{\mathbf{y},\mathbf{z}} \left\{ -\frac{1}{2}\|X^T\mathbf{y} + \nabla^T\mathbf{z}\|_2^2 - \epsilon'\|\mathbf{y}\|_2 - \gamma\|(|\mathbf{z}|)\|_\infty - \mathbf{g}^T\mathbf{y} + \mathbf{f}_{\text{prior}}^T(X^T\mathbf{y} + \nabla^T\mathbf{z}) \right\}, \tag{16}$$

where $\mathbf{z}$ is a spatial-vector image; $|\mathbf{z}|$ is the scalar image produced by taking the vector magnitude of $\mathbf{z}$ at each pixel; the ℓ_∞-norm yields the largest component of the vector argument; and ∇^T is the matrix transpose of ∇. We demonstrate in Sec. III application of CP2-ICTV to both inconsistent and consistent constraint sets. Due to the length of the pseudocode, we present it in the Appendix A, and point out that it can be derived following Ref. 6, using the Moreau identity described in Ref. 10 and an algorithm for projection onto the ℓ_1-ball.[21]

II.F. Summary of proposed convex feasibility methodology

Our previous work in Ref. 6 promoted use of CP Algorithm 1 to prototype convex optimization problems for IIR in CT. Here, we restrict the convex optimization to the form of Eq. (7), allowing the use of the accelerated CP Algorithm 2 with a steeper worst-case convergence rate. Because the proposed optimization Eq. (7) has a generic convex feasibility term, the framework can be regarded as convex feasibility prototyping. The advantage of this approach is twofold: (1) an accelerated CP algorithm is available with an $O(1/N^2)$ convergence rate, and (2) the design of convex feasibility connects better with physical metrics related to the image estimate. To appreciate the latter point, consider the unconstrained counterpart to ICTV. In setting up an objective which is the sum of image TV, data fidelity, and distance from $\mathbf{f}_{\text{prior}}$, two parameters are needed to balance the strength of the three terms. We arrive at

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\{ \frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \alpha_1 \frac{1}{2}\|\mathbf{g} - X\mathbf{f}\|_2^2 + \alpha_2\|\mathbf{f}\|_{\text{TV}} \right\}.$$

As the terms reflect different physical properties of the image, it is not clear at all what values should be selected nor is it clear what the impact of the parameters are on the solution of the unconstrained minimization.

Section III demonstrates use of CP2-EC, CP2-IC, and CP2-ICTV on a breast CT simulation with a limited scanning angular range. The main goals of the numerical examples are to demonstrate use of the proposed convex feasibility framework and convergence properties of the derived algorithms. Even though the algorithms are known to converge within a known worst-case convergence rate, it is still important to observe the convergence of particular image metrics in simulations similar to an actual application.

III. RESULTS: DEMONSTRATION OF THE CONVEX FEASIBILITY ACCELERATED CP ALGORITHMS

We demonstrate the application of the various accelerated CP algorithm instances on simulated CT data generated from the breast phantom shown in Fig. 3. The phantom, described in Refs. 22 and 23, is digitized on a 256 × 256 pixel array. Four tissue types are modeled: the background fat tissue is taken as the reference material and assigned a value of 1.0, the modeled fibro-glandular tissue takes a value of 1.1, the outer skin layer is set to 1.15, and the microcalcifications are assigned values in the range [1.8,2.3]. The simulated CT configuration is described at the beginning of Sec. II.

In the following, the IIR algorithms are demonstrated with ideal data generated by applying the system matrix X to the phantom and with inconsistent data obtained by adding Poisson distributed noise to the ideal data set. We emphasize that the goal of the paper is to address convergence of difficult optimization problems related to IIR in limited angular-range

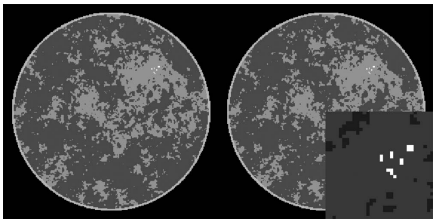


FIG. 3. Breast phantom for the CT limited angular-range scanning simulation. (Left) the phantom in the gray scale window [0.95,1.15]. (Right) the same phantom with a blow-up on the micro-calcification ROI displayed in the gray scale window [0.9,1.8]. The right panel is the reference for all image reconstruction algorithm results.

CT. Thus, we are more interested in establishing that the CP algorithm instances achieve accurate solution to their corresponding optimization problems, and we are less concerned about the image quality of the reconstructed images. In checking convergence in the consistent case, we monitor the conditional primal-dual gap.

For the inconsistent case, we do not have a general criterion for convergence. The conditional primal-dual gap tends to infinity because the dual objective is forced to tend to infinity in order to meet the primal objective, which is necessarily infinity for inconsistent constraints. We hypothesize, however, that CP2-EC minimizes the least-squares problem, Eq. (2), and we can use the gradient magnitude of the least-squares objective to check this hypothesis and test convergence. For CP2-IC, we also hypothesize that it solves the same problem in the inconsistent case, but it is not interesting because we can instead use the parameter-less EC problem. Finally, for CP2-ICTV we do not have a convergence check in the inconsistent case, but we also note that it is difficult to say whether or not a specific instance of ICTV is consistent or not because there are two constraints on quite different image metrics. For this problem, the conditional primal-dual gap is useful for making this determination. If we observe a divergent trend in the conditional primal-dual gap, we can say that the particular choice of TV and data-error constraints are not compatible.

Additionally, we monitor two other metrics as a function of iteration number, the image RMSE is

$$\frac{\|\mathbf{f}-\mathbf{f}_{\text{phantom}}\|_2}{\sqrt{\text{size}(\mathbf{f})}},$$

and the data RMSE is

$$\frac{\|\mathbf{g}-X\mathbf{f}\|_2}{\sqrt{\text{size}(\mathbf{g})}}.$$

We take the former as a surrogate for image quality, keeping in mind the pitfalls in using this metric, see Sec. 14.1.2 of Ref. 24. The latter along with image TV are used to verify that the constraints are being satisfied.

III.A. Ideal data and equality-constrained optimization

We generate ideal data from the breast phantom and apply CP2-EC, with $\mathbf{f}_{\text{prior}}=0$, to investigate its convergence behavior for limited angular-range CT. As the simulations are set up so that X is left-invertible and the data are generated from applying this system matrix to the test phantom, the indicator $\delta_0(X\mathbf{f}-\mathbf{g})$ in Eq. (11) is zero only when $\mathbf{f}$ is the phantom. Observing convergence to the breast phantom as well as the rate of convergence is of main interest here.

In order to have a reference to standard algorithms, we apply linear CG (Ref. 13) and ART to the same problem. Linear CG solves the minimization in Eq. (2), which corresponds to solving the linear system in Eq. (3). The matrix, X^TX, in this equation is symmetric with non-negative singular values. The ART algorithm, which is a form of POCS, solves Eq. (1) directly by cycling through orthogonal projections onto the hyperplanes specified by each row of the linear system.

The results of each algorithm are shown in Fig. 4. As the data are ideal, each algorithm drives the data-error to zero. The linear CG algorithm shows the smallest data RMSE, but we note similar slopes on the log-log plot of CG and CP2-EC during most of the computed iterations except near the end, where the slope of the CG curve steepens. The ART algorithm reveals a convergence slightly faster than CP2-EC, initially, but it is overtaken by CP2-EC near iteration 1000. We also note the impact of the algorithm acceleration afforded by the proposed convex feasibility framework in the comparison of CP2-EC and CP1-EC.

Because X is designed to be left-invertible, we also know that the image estimates must converge to the breast phantom for each of the four algorithms. A similar ordering of the convergence rates is observed in the image RMSE plot, but we note that the values of the image RMSE are all much larger than corresponding values in the data RMSE plots. This stems from the poor conditioning of X, and this point is emphasized in examining the shown image estimates at iteration 10 000 for each algorithm.

While the image RMSE gives a summary metric on the accuracy of the image reconstruction, the displayed images yield more detailed information on the image error incurred by truncating the algorithm iteration. The CP2-EC, CP1-EC, and ART images show wavy artifacts on the left side; the limited-angle scanning arc is over the right-side of the object. But the CG image shows visually accurate image reconstruction at the given gray scale window setting.

This initial result shows promising convergence rates for CP2-EC and that it may be competitive with existing algorithms for solving large, consistent linear systems. But we cannot draw any general conclusions on algorithm convergence, because different simulation conditions may yield different ordering of the convergence rates. Moreover, we have implemented only the basic forms of CG and ART; no attempt at preconditioning CG was made and the relaxation parameter of ART was fixed at 1.

We discuss convergence in detail as it is a major focus of this paper. In Fig. 5, we display the conditional primal-dual gap for the accelerated CP2-EC algorithm compared with use of CP1-EC. First, it is clear that convergence of this gap is slow for this problem due to the ill-conditionedness of X, and we note this slow convergence is in line with the image RMSE curves in Fig. 4. The image RMSE has reached only 10^{-3} after 10^5 iterations. Second, the gap for CP1-EC appears to be lower than that of CP2-EC at the final iteration, but the curve corresponding to CP2-EC went through a similar dip and is returning to a slow downward trend. Third, for a complete convergence check, we must examine the constraints separately from the conditional primal-dual gap, The only constraint in EC is formulated in the indicator $\delta_0(X\mathbf{f}-\mathbf{g})$. In words, this constraint is that the given data and data estimate must be equal or, equivalently, the data RMSE must be zero. We observe in Fig. 4 that the data RMSE is indeed tending to zero. Now that we have a specific example, we reiterate the need for dividing up the convergence check into the conditional primal-dual gap and separate constraint checks. Even though the data RMSE is tending to zero, it is not numerically

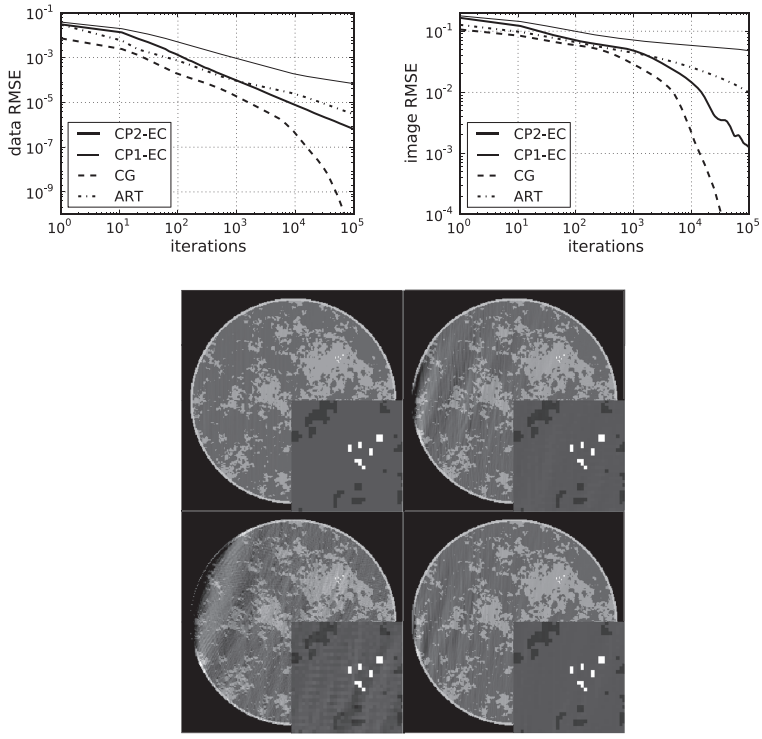


FIG. 4. Results of CP2-EC with ideal, simulated data. Convergence is also compared with CP1-EC, linear CG, and ART. (Top row) (Left) convergence of the four algorithms in terms of data RMSE, and (Right) convergence of the four algorithms in terms of image RMSE. (Bottom row) the image at iteration 10 000 for CG, ART, CP Algorithm 1, and CP2-EC shown in the same gray scale as Fig. 3. The artifacts seen at the right of the images and relatively large image RMSE are indications of the poor conditioning of X. The comparison between CP2-EC and CP1-EC shows quantitatively the impact of the acceleration afforded by CP Algorithm 2.

zero at any iteration and consequently the value of $\delta_0(X\mathbf{f}-\mathbf{g})$ is ∞ at all iterations. Because this indicator is part of the primal objective in Eq. (11), this objective also takes on the value of ∞ at all iterations. As a result, direct computation of the primal-dual gap does not provide a useful convergence check and we need to use the conditional primal-dual gap.

III.B. Noisy, inconsistent data, and equality-constrained optimization

In this section, we repeat the previous simulation with all four algorithms except that the data now contain inconsistency modeling Poisson distributed noise. The level of the noise is selected to simulate what could be seen in a low-dose CT scan. The use of this data model contradicts the application of equality-constrained optimization and EC becomes inconsistent. But nothing prevents us from executing the CP2-EC operations, and accordingly we do so in this subsection. The linear CG algorithm can still be applied in this case, because the optimization in Eq. (2) is well-defined even though there is no $\mathbf{f}$ such that $\mathbf{g}=X\mathbf{f}$. Likewise, the linear system in Eq. (3) does have a solution even when $\mathbf{g}$ is inconsistent. The basic ART algorithm, as with CP2-EC, is not suited to this data model, because it is a solver for Eq. (1), which we know ahead of time has no solution. Again, as with CP2-EC, the steps of ART can still be executed even with inconsistent data, and we show the results here.

In Fig. 6, we show evolution plots of quantities derived from the image estimates from each of the four algorithms. Because the data are inconsistent, the data- and image-error plots have a different behavior than the previous consistent example. In this case, we know that the data RMSE cannot be driven to zero. The algorithms CP2-EC and CG converge on a value greater than zero, while CP1-EC and ART

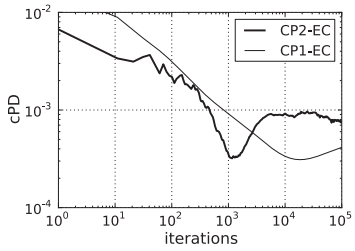


FIG. 5. The conditional primal-dual gap for EC shown for CP2-EC and CP1-EC. This gap is computed by taking the difference between the primal and dual objectives in Eqs. (11) and (12), respectively, after removing the indicator in the primal objective: $cPD = |\frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \frac{1}{2}\|X^T\mathbf{y}\|_2^2 + \mathbf{g}^T\mathbf{y} - \mathbf{f}_{\text{prior}}^T(X^T\mathbf{y})|/\text{size}(\mathbf{f})$. The absolute value is used because the argument can be negative, and we normalize by the number of pixels size($\mathbf{f}$) so that the primal objective takes the form of a mean square error. The prior image $\mathbf{f}_{\text{prior}}$ for this computation is zero. The comparison between CP2-EC and CP1-EC shows quantitatively the impact of the acceleration afforded by CP Algorithm 2.

appear to need more iterations to reach the same data RMSE value.

The image RMSE shows an initial decrease to some minimum value followed by an upward trend. For CG the upward trend begins to level off at 20 000 iterations, while for CP2-EC it appears that this happens near the final 100 000th iteration. For both plots, the results of CP1-EC lag those of the accelerated CP2-EC algorithm.

Turning to convergence checks, we plot the conditional primal-dual gap for EC and the magnitude of the gradient of the least-squares objective from Eq. (2) in Fig. 7. As explained at the beginning of Sec. III, the conditional primal-dual gap tends to infinity for inconsistent convex feasibility problems because the dual objective increases without bound. We observe, in fact, that the conditional primal-dual gap for EC is diverging—a consequence of the inconsistent data used in this simulation. In examining the objective gradient magnitude, the curve for the CG results shows an overall convergence by this metric, because this algorithm is designed to solve the normal equations of the unregularized, least-squares problem in Eq. (2). The ART algorithm shows an initial decay followed by a slow increase. This result is not surprising, because ART is designed to solve Eq. (1) directly and not the least-squares minimization in Eq. (2). As an aside, we point out that in applying ART to inconsistent data it is important to allow the relaxation parameter to decay to zero. Interestingly, CP2-EC and CP1-EC show a monotonic decrease of this gradient.

The resulting gradient magnitude curves indicate convergence of the least-squares minimization, obtained by the CP algorithms. This is surprising, because the conditional primal-dual gap diverges to infinity. Indeed, the magnitude of the dual variable $\mathbf{y}_n$ from the algorithm listed in Fig. 1 increases steadily with iteration number. Even though the dual problem diverges, this simulation indicates convergence of the primal least-squares minimization in that the gradient of this objective is observed to monotonically decrease. There is no proof that we are aware of, which covers this situation, thus we cannot claim that CP2-EC will always converge the least-squares problem. Therefore, in applying CP2-EC in this way it is crucial to evaluate the convergence criterion and to verify that the magnitude of the objective's gradient decays to zero. The conditional primal-dual gap cannot be used as a check for CP2-EC applied to inconsistent data.

The dependence of the gradient magnitude of the unregularized, least-squares objective for the CP2-EC and CG algorithms is quite interesting. Between 10 and 20 000 iterations, CP2-EC shows a steeper decline in this convergence metric. But greater than 20 000 iterations the CG algorithm takes over and this metric drops precipitously. The CG behavior can be understood in realizing that the image has approximately 50 000 unknown pixel values and if there is no numerical error in the calculations, the CG algorithm terminates when the number of iterations equals the number of unknowns. Because numerical error is present, we do not observe exact convergence when the iteration number reaches 50 000, but instead the steep decline in the gradient of the least-squares objective is observed. This comparison between CP2-EC and CG has potential implications for larger systems where the steep drop-off for CG would occur at higher iteration number.

The conditions of this particular simulation are not relevant to practical application because it is already well-known that minimizing unregularized, data-fidelity objectives with noisy data converges to an extremely noisy image particularly for an ill-conditioned system matrix; noting the large values of

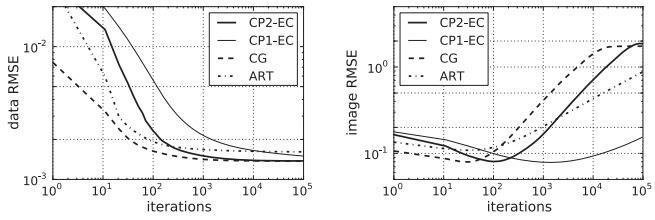


FIG. 6. Metrics of CP2-EC image estimates with noisy and inconsistent, simulated data. Results are compared with CP1-EC, linear CG, and ART. Left, evolution of the four algorithms in terms of data RMSE, and right, evolution of the four algorithms in terms of image RMSE.

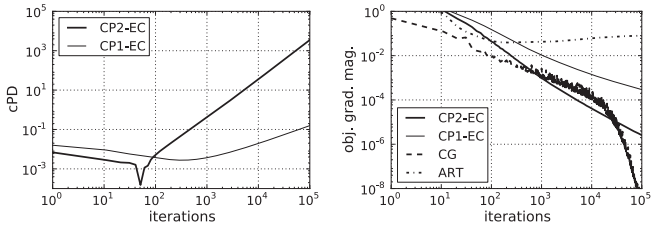


FIG. 7. Convergence plots: the conditional primal-dual gap for EC (left) and the gradient magnitude of the quadratic least-squares objective of Eq. (2) (right). The conditional primal-dual gap is only available for CP2-EC and CP1-EC, while all algorithms can be compared with the objective gradient. The quantity cPD for this problem is explained in the caption of Fig. 5. The convex feasibility problem EC is inconsistent for the simulated noisy data, and as a result cPD diverges to ∞. We hypothesize that CP2-EC converges the least squares minimization Eq. (2), and indeed we note in the gradient plot that CP2-EC yields a decaying objective gradient-magnitude competitive with linear CG and ART. The comparison between CP2-EC and CP1-EC shows quantitatively the impact of the acceleration afforded by CP Algorithm 2.

the image RMSE, we know this to be the case without displaying the image. But this example is interesting in investigating convergence properties. While it is true that monitoring the gradient magnitude of the least-squares objective yields a sense about convergence, we do not know *a priori* what threshold this metric needs to cross before we can say the IIR is converged, see Ref. 19 for further discussion on this point related to IIR in CT. This example in particular highlights the point that an image metric of interest, such as image and data RMSE, needs to be observed to level off in combination with a steady decrease of a convergence metric. For this example, convergence of the image RMSE occurs when the gradient-magnitude of the least-squares objective drops below 10^{-5}, while the data RMSE convergence occurs earlier.

III.C. Noisy, inconsistent data with inequality-constrained optimization

In performing IIR with inconsistent projection data, some form of regularization is generally needed. In using the convex feasibility approach, we apply CP2-IC after deciding on the parameter ϵ'. The parameter ϵ' has a minimum value, below which no images satisfy the data-error constraint, and larger ϵ' leads to greater image regularity. The choice of ϵ' may be guided by properties of the available data or a prior reconstruction. In this case, we have results from Sec. III.B and we note that the data RMSE achieve values below 0.002. Accordingly, for the present simulation we select a tight data-error constraint $\epsilon' = 0.512$, which is equivalent to allowing a data RMSE of $\epsilon = 0.002$. The CP2-IC algorithm selects the image obeying the data-error constraint closest to $\mathbf{f}_{\text{prior}}$, and to illustrate the dependence on $\mathbf{f}_{\text{prior}}$ we present results for two choices: an image of zero values, and an image set to 1 over the support of the phantom. Note that the second choice assumes prior knowledge of the object support and background value of 1. To our knowledge, there is no direct, existing algorithm for solving Eq. (13), and thus we display results for CP2-IC only. One can use a standard algorithm such as linear CG to solve the Lagrangian form of Eq. (13), but this method is indirect because it is not known ahead of time what Lagrange multiplier leads to the desired value of ϵ'.

The results of CP2-IC and CP1-IC are shown in Fig. 8. The data RMSE is seen to converge to the value established by the choice of ϵ'. In the displayed images, there is a clear difference due to the choice of prior image. The image resulting from the zero prior shows a substantial drift of the gray level on the left side of the image. Application of a prior image consisting of constant background values over the object's true support removes this artifact almost completely. These results indicate that use of prior knowledge, when available, can have a large impact on image quality particularly for an ill-conditioned system matrix such as what arises in limited angular-range CT.

Because IC in this case presents a consistent problem, convergence of the CP2-IC algorithm can be checked by the conditional primal-dual gap. This convergence criterion is plotted for CP2-IC and CP1-IC in Fig. 9. The separate constraint check is seen in the data RMSE plot of Fig. 9. We see that the accelerated version of the CP algorithm used in CP2-IC yields much more rapid convergence than CP1-IC. For example, the data RMSE constraint is reached to within 10^{-6} at iteration 1000 for CP2-IC, while this point is not reached for CP1-IC by even iteration 10 000. A similar observation can also be made for the conditional primal-dual gap.

III.D. Noisy, inconsistent data with two-set convex feasibility

For the last demonstration of the convex feasibility approach to IIR for limited-angular range CT, we apply CP2-ICTV, which seeks the image closest to a prior image and respects constraints on image TV and data-error. We are unaware of other algorithms, which address this problem, and only results for CP2-ICTV and CP1-ICTV are shown. In applying CP2-ICTV, we need two constants, ϵ' and γ, and accordingly use of this algorithm is meant to be preceded by an initial image reconstruction in order to have a sense of

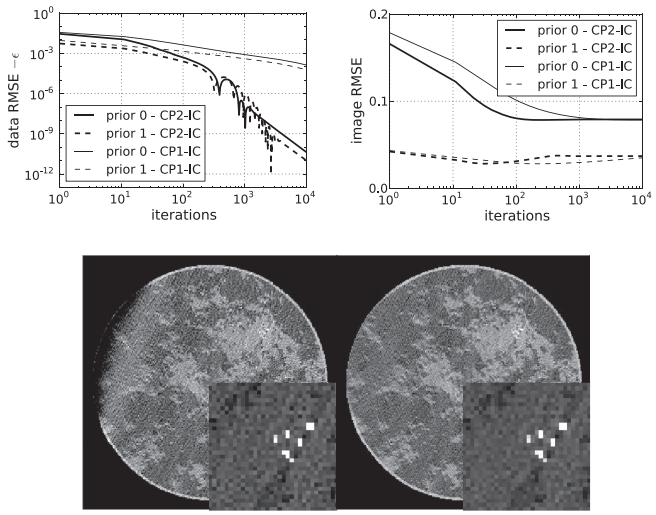


FIG. 8. Results of CP2-IC and CP1-IC with noisy and inconsistent, simulated data. The curves labeled "prior 0" correspond to a zero prior image. The curves labeled "prior 1" correspond to a prior image of 1.0 on the object support. (Top) (left) convergence of the data RMSE to the preset value of $\epsilon = 0.002$ and (right) image RMSE. (Bottom) (Left) "prior 0" final image, and (Right) "prior 1" final image. Gray scales are the same as Fig. 3. The comparison between CP2-EC and CP1-EC shows quantitatively the impact of the acceleration afforded by CP Algorithm 2.

interesting values for the data-error and image TV constraints. From the previous results, we already have information about data-error, and because we have the image estimates, we can also compute image TV values. The image TV values corresponding to the two prior image estimates differ significantly, reflecting the quite different appearance of the resulting images shown in Fig. 8. We follow the use of the support prior

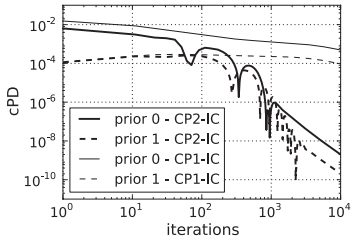


FIG. 9. The conditional primal-dual gap for IC shown for CP2-IC and CP1-IC. This gap is computed by taking the difference between the primal and dual objectives in Eqs. (13) and (14), respectively, after removing the indicator in the primal objective: $\text{cPD} = |\frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 + \frac{1}{2}\|X^T\mathbf{y}\|_2^2 + \epsilon'\|\mathbf{y}\|_2 + \mathbf{g}^T\mathbf{y} - \mathbf{f}_{\text{prior}}^T(X^T\mathbf{y})|/\text{size}(\mathbf{f})$. The absolute value is used because the argument can be negative, and we normalize by the number of pixels size($\mathbf{f}$) so that the primal objective takes the form of a mean square error. The prior image $\mathbf{f}_{\text{prior}}$ for this computation is explained in the text. The comparison between CP2-IC and CP1-IC shows quantitatively the impact of the acceleration afforded by CP Algorithm 2.

image, and take the corresponding value of the image TV of 4400.

In our first example with this two-set convex feasibility problem, we maintain the tight data-error constraint $\epsilon' = 0.512$ (a data RMSE of 0.002) but attempt to find an image with lower TV by selecting $\gamma = 4000$. The results for these constraint set settings, labeled "set 1," are shown in Fig. 10. Interestingly, this set of constraints appears to be just barely infeasible; the CP2-ICTV result converges to an image TV of 4000.012 and a data RMSE of 0.00202. Furthermore, the dual variable magnitude increases steadily, an indication of an infeasible problem. The curves for image TV and data RMSE indicate convergence to the above-mentioned values, but we do not make theoretical claims for convergence of the CP algorithms with inconsistent convex feasibility problems.

In the second example, we loosen the data-error constraint to $\epsilon' = 0.768$ (a data RMSE of 0.0025) and seek an image with lower TV, $\gamma = 3100$, and the results are also shown in Fig. 10. In this case, the constraint values are met by CP2-ICTV, and the resulting image has noticeably less noise than the images with no TV constraint imposed shown in Fig. 8 particularly in the ROI containing the model microcalcifications. The image RMSE for this constraint set in ICTV is 0.029, while the comparable image RMSE from the previous convex feasibility problem, IC, with no TV constraint shown in Fig. 8 is 0.037. Thus we note a drop in image RMSE in adding the image TV constraint, but a true image quality

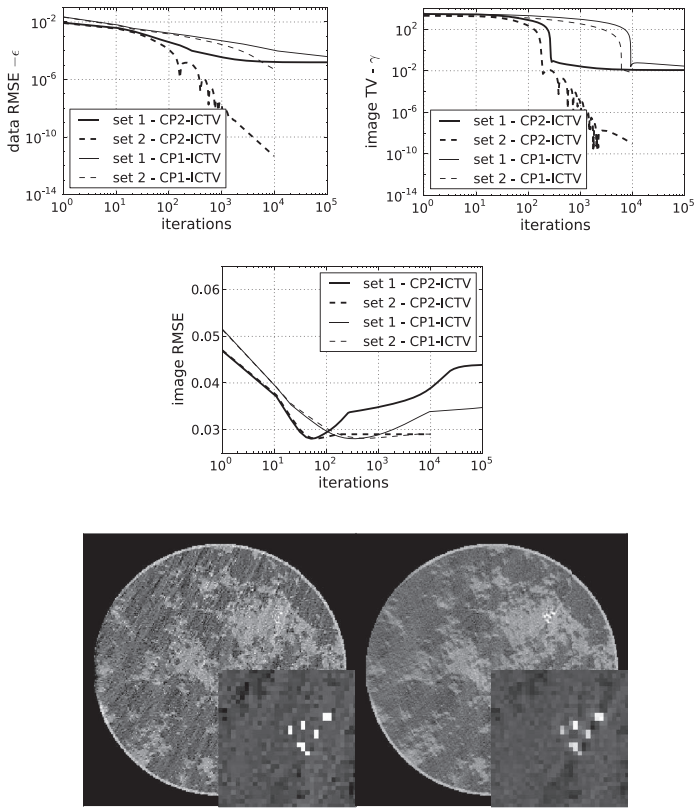


FIG. 10. Results of CP2-ICTV and CP1-ICTV with noisy and inconsistent, simulated data for two different constraint set values: "set 1" refers to choosing $\epsilon' = 0.512$ (a data RMSE of $\epsilon = 0.002$) and $\gamma = 4000$; "set 2" refers to choosing $\epsilon' = 0.768$ (a data RMSE of $\epsilon = 0.0025$) and $\gamma = 3100$. (Top row) (Left) evolution of data RMSE, and (Right) evolution of image TV. (Middle row) evolution of image RMSE. The comparison between CP2-ICTV and CP1-ICTV shows quantitatively the impact of the acceleration afforded by CP Algorithm 2. (Bottom row) (Left) resulting image of "set 1," and (Right) resulting image of "set 2." Gray scales are the same as Fig. 3. Note that the calculation for "set 1" is extended to 10^5 iterations due to slower convergence than the results for "set 2."

comparison would require parameter sweeps in ϵ for IC, and ϵ and γ for ICTV.

Because this constraint set contains feasible solutions, the conditional primal-dual gap can be used as a convergence check for CP2-ICTV. This gap is shown for both sets of constraints in Fig. 11. For CP2-ICTV, there is a stark contrast in behavior between the two constraint sets. The feasible set shows rapid convergence, while the infeasible set shows no decay in the conditional primal-dual gap below 1000 iterations and a steady increase from 1000 to 10 000 iterations. Again, the accelerated CP algorithm used in CP2-ICTV yields a substantially faster convergence rate than CP1-ICTV for this example.

III.E. Comparison of algorithms

With the previous simulations, we have illustrated use of the convex feasibility framework on EC, IC, and ICTV for IIR in CT. The example for EC serves the purpose of demonstrating convergence properties of CP2-EC on the ubiquitous least-squares minimization and establishing that this algorithm has competitive convergence rates with standard algorithms, linear CG, and ART. We do note that CG, on the shown example, does have the fastest convergence rate, but the difference in convergence rate between CP2-EC, CG, and ART is substantially less than their gap with the basic CP1-EC. For convex feasibility problems IC and ICTV, we have

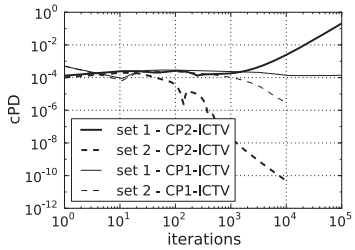


FIG. 11. The conditional primal-dual gap for ICTV shown for CP2-ICTV and CP Algorithm 1. This gap is computed by taking the difference between the primal and dual objectives in Eqs. (15) and (16), respectively, after removing the indicator in the primal objective: cPD $= |\frac{1}{2}\|\mathbf{f}-\mathbf{f}_{\text{prior}}\|_2^2 + \frac{1}{2}\|X^T\mathbf{y}\|_2^2 + \epsilon'\|\mathbf{y}\|_2 + \gamma\|(|\mathbf{z}|)\|_\infty + \mathbf{g}^T\mathbf{y} - \mathbf{f}_{\text{prior}}^T(X^T\mathbf{y} + \nabla^T\mathbf{z})|/\text{size}(\mathbf{f})$. The absolute value is used because the argument can be negative, and we normalize by the number of pixels size($\mathbf{f}$) so that the primal objective takes the form of a mean square error. The prior image $\mathbf{f}_{\text{prior}}$ for this computation is explained in the text. The comparison between CP2-ICTV and CP1-ICTV shows quantitatively the impact of the acceleration afforded by CP Algorithm 2. Note that the calculation for "set 1" is extended to 10^5 iterations due to slower convergence than the results for "set 2."

optimization problems where the current methodology can be easily adapted to solve, but the standard algorithms linear CG and ART cannot easily be applied. Because we have the comparisons of the CP algorithms on the EC simulations and because we have seen convergence competitive with linear CG and ART, we speculate that CP2-IC and CP2-ICTV have competitive convergence rates with any modification of CG or ART that could be applied to IC and ICTV. In short, the convex feasibility framework using CP Algorithm 2 provides a means for prototyping a general class of optimization problems for IIR in CT, while having convergence rates competitive with standard, but more narrowly applicable,

1: $L \leftarrow \|(X,\nabla)\|_2$; $\tau \leftarrow 1$; $\sigma \leftarrow 1/L^2$; $n \leftarrow 0$
2: initialize $\mathbf{f}_0$, $\mathbf{y}_0$, and $\mathbf{z}_0$ to zero vectors
3: $\bar{\mathbf{f}}_0 \leftarrow \mathbf{f}_0$
4: **repeat**
5: $\quad \mathbf{y}'_n \leftarrow \mathbf{y}_n + \sigma(X\bar{\mathbf{f}}_n - \mathbf{g})$; $\mathbf{y}_{n+1} \leftarrow \max(\|\mathbf{y}'_n\|_2 - \sigma\epsilon', 0)\frac{\mathbf{y}'_n}{\|\mathbf{y}'_n\|_2}$
6: $\quad \mathbf{t} \leftarrow \mathbf{z}_n + \sigma\nabla\mathbf{f}_n$
7: $\quad \mathbf{z}_{n+1} \leftarrow \mathbf{t}\left[\left(|\mathbf{t}| - \sigma\,\text{proj}_{\text{Diamond}(\gamma)}(|\mathbf{t}|/\sigma)\right)/|\mathbf{t}|\right]$
8: $\quad \mathbf{f}_{n+1} \leftarrow \left[\mathbf{f}_n - \tau(X^T\mathbf{y}_{n+1} - \mathbf{f}_{\text{prior}} + \nabla^T\mathbf{z}_{n+1})\right]/(1+\tau)$
9: $\quad \theta \leftarrow 1/\sqrt{1+2\tau}$; $\tau \leftarrow \tau\theta$; $\sigma \leftarrow \sigma/\theta$
10: $\quad \bar{\mathbf{f}}_{n+1} \leftarrow \mathbf{f}_{n+1} + \theta(\mathbf{f}_{n+1} - \mathbf{f}_n)$
11: $\quad n \leftarrow n+1$
12: **until** $n \geq N$

FIG. 12. Pseudocode for N steps of the accelerated CP algorithm instance for solving Eq. (15) with parameters ϵ' and γ. Variables are explained in the text, and pseudocode for the function $\text{proj}_{\text{Diamond}(\gamma)}(\mathbf{x})$ is given in Fig. 13.

1: **function** $\text{PROJ}_{\text{DIAMOND}(\gamma)}(\mathbf{x})$
2: $\quad$ **if** $\|\mathbf{x}\|_1 \leq \gamma$ **then**
3: $\quad\quad$ **return** $\mathbf{x}$
4: $\quad$ **end if**
5: $\quad \mathbf{m} = |\mathbf{x}|$
6: $\quad$ Sort $\mathbf{m}$ in descending order: $m_1 \geq m_2 \geq \ldots m_N$
7: $\quad \rho \leftarrow \max j$ such that $m_j - \frac{1}{j}\left(\sum_{k=1}^{j} m_k - \gamma\right) > 0$, for $j \in [1,N]$
8: $\quad \theta \leftarrow (1/\rho)\left(\sum_{k=1}^{\rho} m_k - \gamma\right)$
9: $\quad \mathbf{w} = \max(|\mathbf{x}| - \theta, 0)$
10: $\quad$ **return** $\mathbf{w}\,\text{sign}(\mathbf{x})$
11: **end function**

FIG. 13. Pseudocode for the function $\text{proj}_{\text{Diamond}(\gamma)}(\mathbf{x})$, which projects $\mathbf{x}$ onto the ℓ_1-ball of scale γ. This function appears at Line 7 of algorithm in Fig. 12. The vector $\mathbf{x}$ is taken to be one-dimensional with length N, and the individual components are labeled x_i with index i being an integer in the interval $[1, N]$.

large-scale solvers. Furthermore, concern over algorithm convergence is particularly important for ill-conditioned system models such as those that arise in limited angular-range CT scanning.

Convex feasibility presents a different design framework than unconstrained minimization or mixed optimizations, combining, e.g., data-fidelity objectives with constraints. For example, the field of compressed sensing (CS) (Ref. 25) has centered on devising sparsity exploiting optimization for reduced sampling requirements in a host of imaging applications. For CT, in particular, exploiting gradient magnitude sparsity for IIR has garnered much attention, requiring the solution to constrained, TV-minimization[6,26] or TV-penalized, least-squares.[3–6] The convex feasibility, ICTV, involves the same quantities but can be used only indirectly for a CS-style optimization; the data-error can be fixed and multiple runs with CP2-ICTV for different γ can be performed with the goal of finding the minimum γ given the data and fixed ϵ. On the other hand, due to the fast convergence of CP2-ICTV it may be possible to perform the necessary search over γ faster than use of an algorithm solving constrained, TV-minimization or a combined unconstrained objective. Also, use of ICTV provides direct control over the physical quantities in the optimization, image TV and data-error, contrasting with the use of TV-penalized, least-squares, where there is no clear connection between the smoothing parameter α and the final image TV or data-error. In summary, ICTV provides an alternative design for TV-regularized IIR.

IV. CONCLUSION

We have illustrated three examples of convex feasibility problems for IIR applied to limited angular-range CT, which provide alternative designs to unconstrained or mixed optimization problems formulated for IIR in CT.

One of the motivations of the alternative design is that these convex feasibility problems are amenable to the accelerated CP algorithm, and the resulting CP2-EC, CP2-IC, and CP2-ICTV algorithms solve their respective convex feasibility problems with a favorable convergence rate—an important feature for the ill-conditioned data model corresponding the limited angular-range scan. The competitive convergence rate is demonstrated by comparing convergence of CP2-EC with known algorithms for large-scale optimization. We then note that CP2-IC and CP2-ICTV, for which there is no alternative algorithm that we know of, appears to have similar convergence rates to CP2-EC.

Aside from the issue of convergence rate, algorithm design can benefit from the different point of view offered by convex feasibility. For imaging applications this design approach extends naturally to considering nonconvex feasibility sets,[9,27] which can have some advantage particularly for very sparse data problems. Future work will consider extension of the presented methods to the nonconvex case and application of the present methods to actual data for CT acquired over a limited angular-range scan.

ACKNOWLEDGMENTS

This work is part of the project CSI: Computational Science in Imaging, supported by Grant No. 274-07-0065 from the Danish Research Council for Technology and Production Sciences. This work was supported in part by National Institutes of Health" (NIH) R01 Grant Nos. CA158446, CA120540, and EB000225. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

APPENDIX: PSEUDOCODE FOR CP2-ICTV

The pseudocode for CP2-ICTV appears in Fig. 12, and we explain variables not appearing in Secs. II.E.1 and II.E.2. At Line 6 the symbol ∇ represents a numerical gradient computation, and it is a matrix which applies to an image vector and yields a spatial-vector image, where the vector at each pixel/voxel is either two or three dimensional depending on whether the image reconstruction is being performed in two or three dimensions. Similarly, the variables $\mathbf{t}$ and $\mathbf{z}_n$ are spatial-vector images. At Line 7 the operation "$|\cdot|$" computes the magnitude at each pixel of a spatial-vector image, accepting a spatial-vector image and yielding an scalar image. This operation is used, for example, to compute a gradient-magnitude image from an image gradient. The ratio appearing inside the square brackets of Line 7 is to be understood as a pixelwise division yielding an image vector. It is possible that at some pixels, the numerator and denominator are both zero in which case we define $0/0 = 1$. The quantity in the square brackets evaluates to an image vector, which then multiplies a spatial-vector image; this operation is carried out, again, in pixelwise fashion where the spatial-vector at each pixel of $\mathbf{t}$ is scaled by the corresponding pixel-value. At Line 8, ∇^T is the transpose of the matrix ∇, see Ref. 6 for one possible implementation of ∇ and ∇^T for two dimensions.

The pseudocode for the function $\text{proj}_{\text{Diamond}(\gamma)}(\mathbf{x})$ appears in Fig. 13. This function is essentially the same as what is listed in Fig. 1 of Ref. 21; we include it here for completeness. The "if" statement at Line 2, checks if the input vector $\mathbf{x}$ is already in Diamond(γ). Also, because the function $\text{proj}_{\text{Diamond}(\gamma)}(\mathbf{x})$ is used with a non-negative vector argument in Line 7 of Fig. 12, the multiplication by $\text{sign}(\mathbf{x})$ at the end of the algorithm in Fig. 13 is unnecessary for the present application. But we include this sign factor so that the function applies to any N-dimensional vector.

[a]Electronic mail: sidky@uchicago.edu
[b]Electronic mail: jakj@imm.dtu.dk
[c]Electronic mail: xpan@uchicago.edu

[1]R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography," J. Theor. Biol. **29**, 471–481 (1970).
[2]G. T. Herman, *Image Reconstruction from Projections* (Academic, New York, 1980).
[3]T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, "Implementation of an optimal first-order method for strongly convex total variation regularization," BIT Numer. Math. **52**, 329–356 (2012).
[4]M. Defrise, C. Vanhove, and X. Liu, "An algorithm for total variation regularization in high-dimensional linear problems," Inverse Probl. **27**, 065002 (2011).
[5]S. Ramani and J. Fessler, "A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction," IEEE Trans. Med. Imaging **31**, 677–688 (2012).
[6]E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," Phys. Med. Biol. **57**, 3065–3091 (2012).
[7]P. Combettes, "The convex feasibility problem in image recovery," Adv. Imaging Electron Phys. **95**, 155–270 (1996).
[8]P. L. Combettes, "The foundations of set theoretic estimation," Proc. IEEE **81**, 182–208 (1993).
[9]X. Han, J. Bian, E. L. Ritman, E. Y. Sidky, and X. Pan, "Optimization-based reconstruction of sparse images from few-view projections," Phys. Med. Biol. **57**, 5245–5274 (2012).
[10]A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," J. Math. Imaging Vision **40**, 120–145 (2011).
[11]A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE, New York, 1988).
[12]J. H. Jørgensen, E. Y. Sidky, and X. Pan, "Quantification of admissible undersampling for sparsity-exploiting iterative image reconstruction in X-ray CT," IEEE Trans. Med. Imaging **32**, 460–473 (2013).
[13]J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed. (Springer, New York, 2006).
[14]C. C. Paige and M. A. Saunders, "LSQR: An algorithm for sparse linear equations and sparse least squares," ACM Trans. Math. Softw. **8**, 43–71 (1982).
[15]R. T. Rockafellar, *Convex Analysis* (Princeton University, Princeton, NJ, 1970).
[16]T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *Proceedings of the International Conference on Computer Vision (ICCV 2011)* (IEEE, Barcelona, Spain, 2011), pp. 1762–1769.
[17]Y. Nesterov, "A method of solving a convex programming problem with convergence rate $O(1/k^2)$," Sov. Math. Dokl. **27**(2), 372–376 (1983).
[18]A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," IEEE Trans. Image Process. **18**, 2419–2434 (2009).

[19]J. H. Jørgensen, E. Y. Sidky, and X. Pan, "Ensuring convergence in total-variation-based reconstruction for accurate microcalcification imaging in breast X-ray CT," in *Proceedings of the Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Valencia, Spain, 2011* (IEEE, 2011), pp. 2640–2643.
[20]C. Vogel, *Computational Methods for Inverse Problems* (Society for Industrial Mathematics, Philadelphia, PA, 2002), p. 23.
[21]J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the ℓ_1-ball for learning in high dimensions," in *Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland* (ICML, 2008), pp. 272–279.
[22]J. H. Jørgensen, P. C. Hansen, E. Y. Sidky, I. S. Reiser, and X. Pan, "Toward optimal X-ray flux utilization in breast CT," in *Proceedings of the 11th International Meeting on Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine Potsdam, Germany, 2011* (Fully 3D, 2011), preprint arXiv:1104.1588.
[23]I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," Med. Phys. **37**, 1591–1600 (2010).
[24]H. H. Barrett and K. J. Myers, *Foundations of Image Science* (Wiley, Hoboken, NJ, 2004).
[25]E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," IEEE Signal Process. Mag. **25**, 21–30 (2008).
[26]E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," Phys. Med. Biol. **53**, 4777–4807 (2008).
[27]D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," Inverse Probl. **21**, 37–50 (2005).

Appendix C

Quantitative study of undersampled recoverability for sparse images in computed tomography

---

Submitted to SIAM J. Sci. Comput., 2013.

J. S. Jørgensen, E. Y. Sidky, P. C. Hansen and X. Pan

QUANTITATIVE STUDY OF UNDERSAMPLED RECOVERABILITY FOR SPARSE IMAGES IN COMPUTED TOMOGRAPHY

J. S. JØRGENSEN*, E. Y. SIDKY†, P. C. HANSEN* , AND X. PAN†

Abstract. In x-ray computed tomography (CT) it is generally acknowledged that reconstruction methods that exploit image sparsity allow reconstruction from a significantly reduced number of projections, compared to classical methods. The use of such reconstruction methods is motivated by recent progress in compressed sensing (CS). However, the CS framework provides neither theoretical guarantees of accurate CT reconstruction, nor any relation between sparsity and a sufficient number of measurements for recovery, i.e., perfect reconstruction from noise-free data. We consider reconstruction through 1-norm minimization, as proposed in CS, from data obtained using a standard fanbeam sampling pattern in CT, i.e., no artificial random sampling patterns as is common in CS. Through computer simulations we establish quantitatively a relation between the image sparsity and the sufficient number of measurements for recovery. To do this, we develop a so-called relative sparsity-sampling diagram for empirically studying this relation over random realizations of sparse test images from parametrized image classes. Our main result is the observation that the transition from non-recovery to recovery is sharp in the sense that same-sparsity phantom realizations require essentially the same number of CT projections for recovery. We show that the specific behavior depends on the type of image, and that the same quantitative relation holds independently of image size and is robust to small amounts of additive Gaussian noise.

Key words. Inverse Problems, Computed Tomography, Image Reconstruction, Compressed Sensing, Sparse Solutions

AMS subject classifications. 90C90, 15A29, 94A08, 44A12

1. Introduction.

In x-ray computed tomography (CT) an image of an object is reconstructed from projections obtained by measuring the attenuation of x-rays passed through the object. Motivated by the need to reduce the exposure to radiation, there is a growing interest in low-dose CT, cf. [30] and references therein. This is relevant in medical imaging to reduce the risk of radiation-induced cancer, and in biomedical imaging where high doses can damage the specimen under study.

Classical reconstruction methods are based on closed-form analytical or approximate inverses of the continuous forward operator; examples are the filtered backprojection method [20] and the Feldkamp-Davis-Kress method for cone-beam CT [13]. Their main advantages are low memory demands and computational efficiency, which make them the current methods of choice in commercial CT scanners [3, 21]. However, they are known to have limitations on reduced data.

One alternative is to use an algebraic formulation where the forward operator is fully discretized, leading to a large sparse system of linear equations. This approach can handle geometries for which no analytical inverse is available, such as non-standard scanning geometries. Another advantage is that the data reduction arising from low-dose imaging can sometimes be compensated for by incorporating prior information about the image. This is possible through a variational formulation in which the objective function reflects the desired image properties, such as smoothness, non-negativity, or as in our case sparsity, i.e., having a representation with few non-zero coefficients.

*Department of Applied Mathematics and Computer Science, Technical University of Denmark, Matematiktorvet, Building 303B, 2800 Kgs. Lyngby, Denmark ({jakj,pcha}@dtu.dk).

†Department of Radiology, University of Chicago, 5841 S. Maryland Ave., Chicago, IL 60637, USA ({sidky,xpan}@uchicago.edu).

Developments in compressed sensing (CS) [6, 9] show potential for a reduction in data while maintaining or even improving reconstruction quality. This is made possible by exploiting image sparsity; loosely speaking, if the image is "sparse enough", then it admits accurate reconstruction from undersampled data. We refer to such methods as *sparsity-exploiting methods*.

Different types of sparsity can be relevant in CT. In reconstruction of blood vessels [18] the image itself can be considered sparse. In reconstruction of cross-sections of the human body, which consists of well-separated areas of relatively homogeneous tissue, the image gradient is approximately sparse, and this property can be exploited in a reconstruction algorithm by minimizing the total variation (TV) of the image [27]. Studies using simulated as well as clinical data have demonstrated that sparsity-exploiting methods indeed allow for reconstruction from fewer projections [4, 15, 26, 28, 29]. In spite of these positive results, we still lack a fundamental understanding of conditions — especially in a quantitative sense — under which such methods can be expected to perform well in CT.

The present paper investigates the possible relation between the image sparsity and the sufficient number of CT views for accurate reconstruction of the image. To simplify our analysis we focus on images with sparsity in the image domain and reconstruction through minimization of the image 1-norm subject to a data equality constraint, as motivated by CS. These studies are interesting in their own right and they set the stage for forthcoming studies of other regularizers, such as TV, as well as other types of sparsity.

We are unaware of theoretical results from CS that cover the mathematical model for CT. Instead — inspired by work of Donoho and Tanner [8] — we can use computer simulations to empirically study *recoverability* within well-defined classes of sparse images. Specifically, we are interested in the average number of projections sufficient for exact recovery of an image as function of the image sparsity. An advantage of this approach (instead of relying on specific theoretical results for unnatural sampling patterns) is that it can be easily extended to systems of increasing levels of realism.

CS image recovery typically rely on random sampling, but it is widely acknowledged in the CT community that undersampled image recovery can be achieved with structured sampling patterns used in commercial CT scanners. However, we still lack a quantitative understanding of which factors influence the reconstruction quality. Here we address this shortcoming by establishing the following empirical results:

1. There is a quantitative relation between image sparsity and sufficient sampling for recovery.
2. There is a sharp transition from non-recovery to recovery.
3. The specific relation varies with respect to different image classes.
4. The relation holds independently of the image dimension.
5. The relation appears to be robust with respect to additive Gaussian noise.

Another interesting result is that the advantage of using a sparsity-exploiting method is significant, even for images with relatively many non-zeros.

We believe that our findings shed light on the connection between sparsity and sufficient sampling in CT, and maybe more importantly, provide an operational tool of use for CT engineers designing a CT system based on exploiting image sparsity.

Our paper is organized as follows. Section 2 gives the problem formulation and the sparsity-exploiting reconstruction method, and it introduces the concept of recoverability in CT. Section 3 describes our numerical simulations, including details of the CT imaging model, generation of sparse phantoms, and how to robustly solve the

reconstruction optimization problems. Section 4 presents an overview of our results, and we conclude with a discussion in Section 5.

2. Sparsity-exploiting reconstruction methods. The purpose of this brief section is to define the notation, the algebraic formulation, and the reconstruction method used throughout the study.

2.1. Constrained algebraic reconstruction. We consider the discrete inverse problem of recovering a signal $x_{\text{orig}} \in \mathbb{R}^N$ from (usually noisy) measurements $b \in \mathbb{R}^M$. The imaging model, which is assumed to be linear and discrete-to-discrete [2], relates the image and the data through a *system matrix* $A \in \mathbb{R}^{M\times N}$,

$$b = A\,x + e, \tag{2.1}$$

where the elements of $x \in \mathbb{R}^N$ are pixel values stacked into a vector and $e \in \mathbb{R}^M$ represents additive noise. To solve (2.1) it is often necessary to impose regularization in order to reduce noise amplification in the inversion as well as to restrict the set of solutions in case of an underdetermined and/or rank deficient problem. A common type of regularization takes the form: $\min_x\ J(x)$ s.t. $\|A\,x - b\|_2 \le \eta$, where $J(x)$ is a *regularizer*, i.e., a function selected to impose some condition on the image that reflects prior knowledge or assumptions. In this work we use $J(x) = \|x\|_1$, i.e.,

$$\textbf{L1}_\eta\text{:} \qquad \min_x\ \|x\|_1 \qquad \text{s.t.} \qquad \|Ax - b\|_2 \le \eta, \tag{2.2}$$

which is known to often produce a sparse x, as discussed below. The regularization parameter η reflects the amount of noise in the data, and in the limit $\epsilon \to 0$ we obtain the equality-constrained problem:

$$\textbf{L1}\text{:} \qquad \min_x\ \|x\|_1 \qquad \text{s.t.} \qquad A\,x = b. \tag{2.3}$$

The corresponding problem

$$\textbf{L2}\text{:} \qquad \min_x\ \|x\|_2 \qquad \text{s.t.} \qquad A\,x = b, \tag{2.4}$$

which arises from Tikhonov regularization, gives the unique minimum-norm solution, i.e., the vector of minimal 2-norm among all vectors satisfying $A\,x = b$.

The inequality-constrained problem is of more practical interest than (2.3) because it allows for noisy and inconsistent measurements, but its solution depends in a complex way on the noise and inconsistencies in the data, as well as the choice of the parameter η. Studies of the equality-constrained problem (2.3), on the other hand, provide an basic understanding of a given regularizer's reconstruction potential, independent of specific noise. Therefore, we focus in the present study on the ideal equality-constrained problem. This means that we do not consider uncertainties in the system matrix or the influence of the regularization parameter. We do, however, study the robustness of the results with respect to additive Gaussian noise.

2.2. Recoverability of problem instances. The interest in **L1** (as well as TV and other sparsity-exploiting methods) is motivated by recent developments in CS demonstrating that it is possible to recover x_{orig} from a reduced number of measurements [6, 9]. The underlying assumption is that the image has few non-zero pixels, or that it is sparse in some other representation (such as after applying a discrete gradient transformation to the image). We call a vector $x \in \mathbb{R}^N$ with k non-zero elements k-sparse and we define the

$$\textbf{relative sparsity:} \qquad \kappa = k/N. \tag{2.5}$$

Moreover, we call a tuple (x_{orig}, A, $b = A\,x_{\text{orig}}$) a *problem instance* and say that it is *recoverable* if solving **L1** produces a solution $x^\star$ that is identical to x_{orig}. Existence and uniqueness will be discussed in Section 3.4.

There has been much work on establishing results stating for which matrices A problem **L1** is capable of recovering the sparsest solution, cf. [5] and references therein. We give an example of such a theorem based on the mutual coherence μ_A of a matrix A: For a full-rank $A \in \mathbb{R}^{M\times N}$ with $M < N$, if a k-sparse solution x to $A\,x = b$ satisfies

$$k < \frac{1}{2}\left(1 + \frac{1}{\mu_A}\right), \qquad \mu_A = \max_{1\le i,j\le N} \frac{|a_i^T a_j|}{\|a_i\|_2 \cdot \|a_j\|_2}$$

(where a_ℓ is the ℓth column of A), then it is the unique solution to **L1**; a smaller μ_A leads to a larger bound on the sparsity of a signal that is guaranteed to be recovered. Many theoretical recovery results exist, most notably relying on the so-called "spark" [10] and restricted isometry property [7] of a matrix. While some results are deterministic, many of them are probabilistic in the sense that if the elements of A are selected at random from certain probability distributions, then with "overwhelming probability" **L1** will recover the original signal [6].

2.3. Application to CT. It is generally accepted that these theoretical results are of limited practical use [12], since the requirements are generally NP-hard to check [23], and/or they provide very pessimistic bounds on the sparsity of signals that can be recovered. Better results are available for certain special matrices, such as those with elements drawn from a Gaussian distribution, but those results do not carry over to the system matrices encountered in CT. For example, the CT matrices considered in the present work have mutual coherences between 0.5 and 1, leading to guaranteed recovery only of images with a single non-zero element.

Instead, recoverability can be studied empirically. Such studies have been conducted for certain practical systems (see, e.g., [1, 22]) but we are unaware of any systematic recoverability studies specifically for CT system matrices. Our empirical study is inspired by the work by Donoho and Tanner (DT) [8] who studied empirical recoverability using a *phase diagram* of the (δ, ρ)-plane, where:

$$\textbf{undersampling fraction:}\quad \delta = M/N, \qquad \textbf{sparsity fraction:}\quad \rho = k/M. \tag{2.6}$$

For certain classes of randomly generated matrices and test images, DT were able to prove existence of a sharp transition from non-recovery to recovery, and verify the result in empirical studies. Although we do not derive similar theoretical results for CT matrices, we can still conduct similar empirical recoverability studies using the DT phase diagram and the related relative sparsity-sampling (RSS) diagram, which we introduce in Section 4. Since our interest is recoverability when using a specific deterministic CT matrix, we do not use ensembles of randomly generated matrices; only the test images are randomly generated.

We note that it is possible to construct examples of k-sparse vectors for small k that cannot be recovered from CT measurements [11, 24], implying that we cannot hope to obtain useful results on guaranteed recovery of *all* k-sparse images. However, these constructed vectors might be pathological and very different from actual images occurring in practice, and for this reason we are more interested in determining *average-case* recovery results for specific classes of images.

We will empirically establish a quantitative relation between the number of measurements and the sparsity of x_{orig} sufficient for recoverability. In order to do that,

we conduct randomized simulations where we generate ensembles of images of varying sparsity and CT system matrices corresponding to varying number of projections; then we use **L1** for reconstruction, and we check for recovery. Since different phantoms of same class and sparsity might require different number of views to be recovered, we are interested in the average-case recovery over the phantom ensembles.

3. Experiment design. In this section we describe the test problems used in our studies, as well as our approach to solving the reconstruction problem numerically.

3.1. CT imaging geometry. There is no generic CT reconstruction model; the geometry depends on the scanner type, and in the reconstruction algorithm one can adjust the number of projections and the number of pixels (to trade-off resolution for signal-to-noise ratio), re-bin or interpolate the data to obtain additional "data", use other basis functions than pixels, etc.

As a specific example we consider a 2D fan-beam geometry with equi-angular views. We consider a square domain of $N_{\text{side}} \times N_{\text{side}}$ pixels, and due to rotational symmetry we restrict the region-of-interest to be within a circular mask inside the square domain consisting of approximately $N = \lceil \pi/4 \cdot N_{\text{side}}^2 \rceil$ pixels. The source-to-center distance is $2N_{\text{side}}$, and the fan-angle $28.07°$ is set to precisely illuminate the circular mask. The number of views (or projections) is denoted N_{v}. The rotating detector is assumed to consist of $2N_{\text{side}}$ bins, so the total number of measurements is $M = 2N_{\text{side}}N_{\text{v}}$. The $M \times N$ system matrix A is computed by means of the MATLAB package AIR Tools [16].

3.2. Sparse phantom classes. By a *class of phantoms* we mean a set of phantom images described by a set of specifications, such that we can generate random realizations from the class. We refer to such an image as a *phantom instance* from the class, and multiple phantom instances from the same class form a *phantom ensemble*.

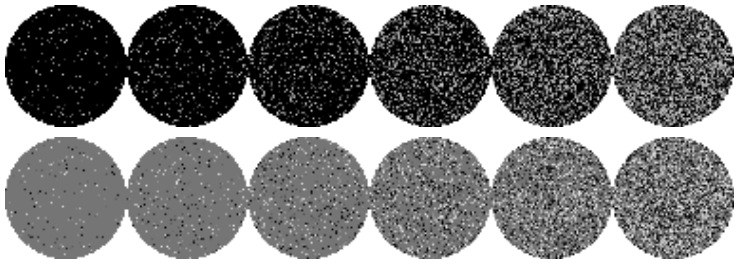


FIG. 3.1. *Sparse phantom image instances of class* **spikes** *(1st row, black is 0, white is 1), and of class* **signedspikes** *(2nd row, black is −1, white is 1). Relative sparsity from left to right is* $\kappa = 0.05, 0.10, 0.20, 0.40, 0.60, 0.80$.

For the **spikes** class, given an image size N and a target relative sparsity κ, we generate a phantom instance as follows: starting from the zero image, randomly select $k = \text{round}(\kappa N)$ pixel indices, and set each selected pixel in $x \in \mathbb{R}^N$ to a random number from a uniform distribution over $[0, 1]$. Figure 3.1 shows examples of **spikes** phantom instances for varying κ. This class is deliberately designed to be as simple as possible and it does not mimic any particular application; it is solely used to study the generic case of having a sparse image.

The **signedspikes** class is essentially identical to the **spikes** class, except the non-zero pixel values are uniformly distributed over $[-1, 1]$; see Fig. 3.1. In standard CT reconstruction the attenuation coefficient is always non-negative, but in certain applications a background attenuation is subtracted, thereby causing x to have both positive and negative values. As we will show in Section 4, the modification to allow negative pixel values leads to a considerable change in recoverability.

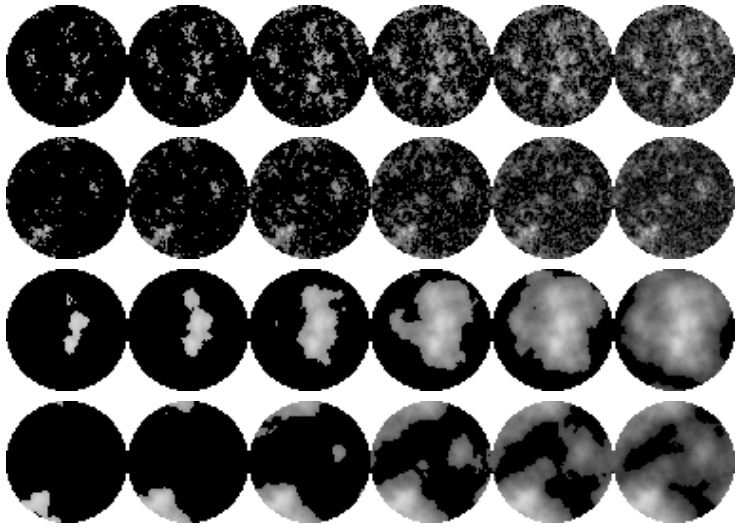


FIG. 3.2. *Two sets of phantom image instances of class* **1-power** *(1st and 2nd rows, black is 0), and of class* **2-power** *(3rd and 4th rows, black is* 0*). Relative sparsities from left to right are* $\kappa = 0.05, 0.10, 0.20, 0.40, 0.60, 0.80$.

The ***p*-power** class models a more realistic type of images, namely background tissue in the female breast. The idea is to introduce structure to the pattern of non-zero pixels by creating correlation between neighboring pixels. Our procedure is based on [25] followed by thresholding to obtain many pixel values strictly equal to 0; the amount of structure is governed by a parameter p:

1. Create an $N_{\text{side}} \times N_{\text{side}}$ phase image P with values drawn from a Gaussian distribution with zero mean and unit standard deviation.
2. Create an $N_{\text{side}} \times N_{\text{side}}$ amplitude image U with pixel values

$$U(i,j) = \left(\left(\frac{2i-1}{N} - 1 \right)^2 + \left(\frac{2j-1}{N} - 1 \right)^2 \right)^{-p/2}, \quad i,j = 1, \ldots, N_{\text{side}}.$$

3. For all pixels (i, j) compute $F(i,j) = U(i,j)e^{2\pi \hat{\imath} P(i,j)}$, with $\hat{\imath} = \sqrt{-1}$.
4. Compute the magnitude of the 2D inverse discrete Fourier transform of F.
5. Restrict this square image to the disk-shaped mask.
6. Keep the $k = \text{round}(\kappa N)$ largest pixel values and set the rest to 0.

Figure 3.2 shows examples of phantoms from classes **1-power** and **2-power**. Both have more structure than the **spikes** phantoms, and the structure increases with p.

We do not claim that our image classes are entirely realistic models, e.g., of breast tissue (although ***p*-power** was developed with this application in mind). Our goal is to study simplified *parametrized* sparse images, and we find that the choice of class has only limited influence on recovery results. Hence, we find it unlikely that an even more realistic phantom class will produce significantly different results.

3.3. Robust solution of optimization problems. Given a numerically computed solution, the robustness of the decision regarding recovery depends on the accuracy of the solution. False conclusions may result from incorrect or inaccurate solutions. To robustly solve the optimization problems **L1** and **L2** we must therefore use a numerical method which gives a clear indication of whether a correct solution, within a given accuracy, has been computed. Our choice is the package MOSEK [19], which uses a primal-dual interior-point method. MOSEK is equipped with numerous sophisticated features to handle numerical difficulties, and it issues warnings when it fails to compute an accurate solution. In all problem instances considered in our simulation studies, MOSEK managed to return a certified accurate solution.

To solve **L2** with MOSEK we recast it as a quadratic program, which is readily solved by MOSEK. To solve **L1** using MOSEK we recast it as the linear program $\min_w \mathbf{1}^T w$ s.t. $A\,x = b$ and $-w \le x \le w$, $\mathbf{1} \in \mathbb{R}^N$.

3.4. Simulations. Given the imaging model, the method for generating sparse phantom images, and a robust optimization algorithm, we are in the position to carry out randomized simulation studies of recoverability within a phantom class, as a function of relative sparsity and number of views. The design of a single basic simulation consists of the following steps: 1) Generate an instance $(A, x_{\text{orig}}, b = Ax_{\text{orig}})$, 2) solve **L1** numerically to obtain $x^\star$, and 3) numerically test for recovery using

$$\frac{\|x^\star - x_{\text{orig}}\|_2}{\|x_{\text{orig}}\|_2} < \epsilon, \tag{3.1}$$

where the threshold ϵ is chosen based on the accuracy of the optimization algorithm; empirically we found $\epsilon = 10^{-4}$ to be well-suited in our set-up.

While **L2** has a unique solution because $\|x\|_2^2$ is strictly convex, it is not necessarily the case for **L1**. For the test problems considered here, existence of a solution is guaranteed by the way we generate data, while uniqueness can not be known in advance. The solution set may consist either of a single image or an entire hyperface or hyperedge on the 1-norm ball. The particular solution found depends on the optimization algorithm, and therefore our conclusions of recoverability by **L1** are subject to our use of MOSEK. We do not specifically check for uniqueness, however, in the event of infinitely many solutions, it is unlikely that any optimization algorithm will select precisely the original image, so we believe that our observations of recoverability correspond to existence of a unique solution.

4. Simulation results. We start by introducing some notation that is useful for the following discussion. For a given problem instance, the *sufficient view number* $N_{\text{v}}^{\text{suf}}$ denotes the smallest number of views that causes A to have full column rank. The **L1** *recovery view number* $N_{\text{v}}^{\mathbf{L1}}$ denotes the smallest number of views for which recovery is observed for all $N_{\text{v}} \ge N_{\text{v}}^{\mathbf{L1}}$. Using $N_{\text{v}}^{\text{suf}}$ as a reference point for full sampling, we define two useful quantities:

$$\textbf{relative sampling:} \qquad \mu = N_{\text{v}}/N_{\text{v}}^{\text{suf}}, \tag{4.1}$$

$$\textbf{relative sampling for recovery:} \qquad \mu^{\mathbf{L1}} = N_{\text{v}}^{\mathbf{L1}}/N_{\text{v}}^{\text{suf}}, \tag{4.2}$$

cf. the relative sparsity κ defined in (2.5).

One of our main contributions is the so-called relative sparsity-sampling (RSS) diagram, introduced in Section 4.3, in which the recovery percentage over an ensemble of phantoms is plotted as function of the relative sparsity and the relative sampling. The RSS diagram reveals a sharp transition from non-recovery to recovery and a monotonically increasing relation between relative sparsity and relative sampling for recovery. The subsequent sections study how the RSS diagrams change with variations such as to the image size, phantom class and addition of noise to the data.

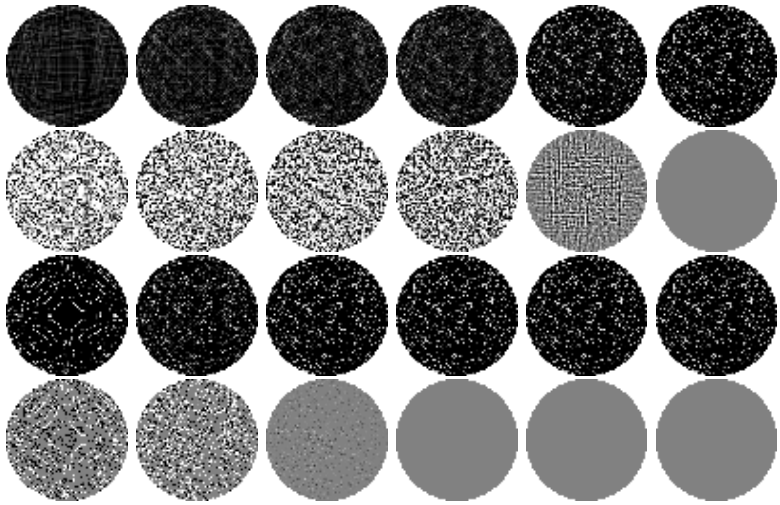


FIG. 4.1. *Reconstructions of the* **spikes** *phantom with* $N_{\text{side}} = 64$, *relative sparsity* $\kappa = 0.20$. *1st row:* **L2** *reconstructions (black is* 0, *white is* 1*). 2nd row:* **L2** *minus original image (black is* -0.1, *white is* 0.1*). 3rd row:* **L1** *reconstructions (black is* 0, *white is* 1*). 4th row:* **L1** *minus original image (black is* -0.1, *white is* 0.1*). Columns:* 4, 8, 10, 12, 24 *and* 26 *views.*

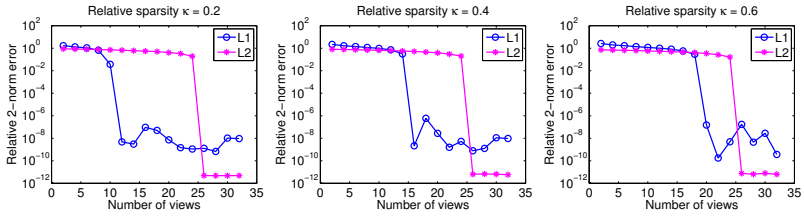


FIG. 4.2. *Numerical recovery measure from* (3.1) *vs. view numbers for* **L1** *and* **L2** *reconstruction of* **spikes** *phantoms with relative sparsity values* $\kappa = 0.2$, 0.4 *and* 0.6. *The numerical accuracy of the optimization algorithm is reflected in the attained non-zero errors.*

4.1. Recovery from undersampled data. We first establish that **L1** is capable of recovering an image from undersampled CT measurements, and we compare with the **L2** reconstruction. We use a phantom x_{orig} from the **spikes** class with

$N_{\text{side}} = 64$, leading to $N = 3228$ pixels inside the mask. The relative sparsity is set to $\kappa = 0.20$, which yields 646 non-zeros. We consider reconstruction from data corresponding to $2, 4, 6, \ldots, 32$ views; the smallest and largest system matrices are of sizes 256×3228 and 4096×3228, respectively. Selected reconstructed images $x^\star$ from solving **L1** and **L2** are shown in Figure 4.1 along with the error images $x^\star - x_{\text{orig}}$ to better visualize the sudden drop in error when the image is recovered.

Recall that the minimum-norm **L2** solution is typically non-sparse [12]. Therefore, we expect to need a full-rank system matrix for **L2** to recover the original sparse image, and our results confirm this: the **L2** reconstructions gradually improve with more views but recovery is not observed until $N_{\text{v}} = N_{\text{v}}^{\text{suf}} = 26$. At $N_{\text{v}} = 24$, the matrix is 3072×3228 and $\text{rank}(A) = 3052$, and the minimum-norm **L2** solution is not the original. At $N_{\text{v}} = 26$, the matrix is 3328×3228 and full-rank, so x_{orig} is recovered.

For **L1**, recovery occurs already at $N_{\text{v}} = N_{\text{v}}^{\mathbf{L1}} = 12$, where A has size 1536×3228 and rank 1524. Evidently, in spite of the large null space of A, in this case, **L1** selects the original image. When N_{v} increases, **L1** continues to recover the original up to and beyond $N_{\text{v}} = N_{\text{v}}^{\text{suf}} = 26$, where the matrix becomes full-rank. Figure 4.1 thus demonstrates the well-known potential of **L1** recovery; here in the setting of undersampled CT measurements.

In order to investigate quantitatively a possible relation between image sparsity and sufficient sampling for **L1** recovery we repeat the $\kappa = 0.2$ experiment for $\kappa = 0.4$ and 0.6. Figure 4.2 shows the numerical recovery measures from (3.1) for **L1** and **L2** as a function of N_{v}. For **L2** the behavior is independent of the relative sparsity, as expected. For **L1**, on the other hand, $N_{\text{v}}^{\mathbf{L1}}$ takes the values 12, 16 and 20, indicating a very simple relation between sparsity and **L1** recovery view number.

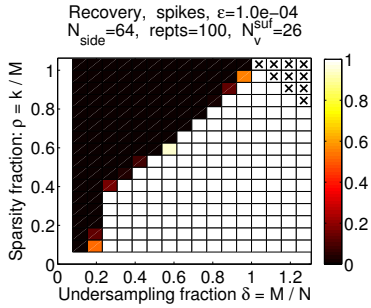


FIG. 4.3. *DT phase diagram for recovery of* **spikes** *phantoms,* $N_{\text{side}} = 64$.

4.2. Recoverability studies using the DT diagram. In general, we can not expect all phantom instances of the same relative sparsity to have the same $N_{\text{v}}^{\mathbf{L1}}$, and in fact we observe some variation. One way to study this variation is through the DT phase diagram described in Section 2.2. For the **spikes** class with $N_{\text{side}} = 64$ we consider reconstruction with undersampling fraction $\delta = M/N$ for $M = 2N_{\text{side}}N_{\text{v}}$ and $N_{\text{v}} = 2, 4, 6, \ldots, 32$. At each N_{v} we consider sparsity fractions $\rho = k/M = 1/16, 2/16, \ldots, 16/16$. We test for recovery by reconstruction using the same system matrix A for 100 phantom instances at each (δ, ρ).

The resulting diagram with the percentage of instances recovered at each (δ, ρ) is

shown in Figure 4.3. At very low sampling ($N_v = 2$ and 4, i.e., $\delta = 0.08$ and 0.16) no phantoms can be recovered. An important observation is the very sharp phase transition from non-recovery to recovery, meaning that the variation of $N_v^{\mathbf{L1}}$ within the same phantom class is very limited. To the best of our knowledge, this analysis has not been done for CT-matrices before, and we therefore believe the sharp transition to be a new observation.

Normalizing the sparsity by the number of measurements causes a small problem for the diagram. When $\delta > 1$, values of ρ close to 1 can lead to $k > N$, for which it is impossible to construct any instances; those cases are shown with $\times$-symbols.

4.3. The relative sparsity-sampling (RSS) diagram. Our main interest is the relation between the relative sparsity $\kappa = k/N$ and the relative sampling for recovery $\mu^{\mathbf{L1}} = N_v^{\mathbf{L1}}/N_v^{\text{suf}}$. For our purpose, we find interpretation of the DT diagram inconvenient, as it uses different quantities on the axes. Instead, we choose to visualize the percentage of instances recovered in the (κ, μ)-plane. We refer to such a diagram as a *relative sparsity-sampling* (RSS) diagram. Figure 4.4 (left) shows the RSS diagram corresponding to the DT phase diagrams in Figure 4.3, created by reconstructing 100 **spikes** phantoms for $\kappa = 0.025, 0.05, 0.1, 0.2, \ldots, 0.9$ and $N_v = 2, 4, 6, \ldots, 32$. Figure 4.4 (right) shows the *average* $\mu^{\mathbf{L1}}$ over all instances at each κ. In order to quantify the possible deviation of the empirical average compared to the true unknown mean we also show the 99% confidence interval estimated using the empirical standard deviation, illustrated by small horizontal lines.

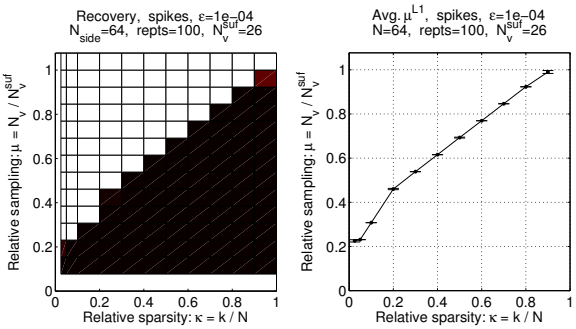


FIG. 4.4. *Left: Relative sparsity-sampling (RSS) diagram for recovery of the* **spikes** *class at* $N_{\text{side}} = 64$*; the colorscale is as in Figure 4.3. Right: Average relative sampling and its* 99% *confidence interval for recovery over the phantom instances.*

As in the DT phase diagram in Figure 4.3, the RSS diagram reveals a sharp transition from non-recovery to recovery meaning that the variation of $N_v^{\mathbf{L1}}$ over phantom instances is almost negligible. The 99% confidence intervals are very narrow, in fact, in several cases of width zero, due to zero variation of $N_v^{\mathbf{L1}}$, which agrees well with the visual observation of a sharp transition. The relative sampling for recovery $\mu^{\mathbf{L1}}$ increases monotonically with the relative sparsity κ, although not in a linear way. As $\kappa \to 0$, the relative sampling for recovery $\mu^{\mathbf{L1}}$ also approaches 0, and similarly $\mu^{\mathbf{L1}} \to 1$ for $\kappa \to 1$, confirming that when the image is no longer sparse, **L1** gives no advantage over **L2**.

The RSS diagram also gives *quantitative* information on the recovery view number for **L1**. Assume, for example, that we are given an image of relative sparsity $\kappa = 0.1$,

how many views would suffice for recovery? The RSS diagram shows that at $\kappa = 0.1$, we have $\mu^{\mathbf{L1}} = 0.31$, which corresponds to $N_{\mathrm{v}}^{\mathbf{L1}} = 8$ views. If the phantom has $\kappa = 0.6$, we obtain $\mu^{\mathbf{L1}} = 0.77$ and $N_{\mathrm{v}}^{\mathbf{L1}} = 20$ views. Note that the RSS-diagram works equally well for answering the opposite question, namely, what is the maximal relative sparsity that, on average, will allow recovery from 20 views?

We emphasize that our goal is not to advocate **L1** as the "best" optimization problem for undersampled CT reconstruction but to propose the RSS diagram as a tool systematically comparing variations of the optimization problem. For example, the **spikes** class has non-negative entries so a natural idea is to impose a non-negativitity constraint in the hope of achieving accurate reconstruction at even lower relative sampling. We implemented this idea and constructed the RSS diagram, which turned out to be identical to the diagram for **L1**. From this we conclude that the limiting factor for reducing the relative sampling for recovery is not negative solution elements, because the **L1** reconstruction is, in this case, already negative without incorporating the constraint.

Clearly, the RSS diagram introduced here is strongly inspired by the DT phase diagram, but for several reasons we find the RSS diagram more intuitive to interpret for our CT applications:

1. A definition of the undersampling fraction relative to N, as used in the DT phase diagram, only makes sense when $M = N$ yields a full-rank matrix. This is not necessarily the case in CT, so a slightly different measure of undersampling, $\mu = N_{\mathrm{v}}/N_{\mathrm{v}}^{\mathrm{suf}}$, is required. (For the CT geometry studied here, $\delta \approx \mu$ because the N_{v} that yields full rank is close to $M = N$; but to be precise we still make the distinction.)
2. The sparsity fraction ρ for the DT phase diagram is relative to the number of measurements M, whereas our sparsity κ is relative to the number of pixels N. In the DT phase diagram, our relative sparsity is constant along hyperbolic curves instead of straight lines; to see this, note that a constant $\kappa = c$ is equivalent to $\rho\delta = c$, which is a hyperbola in the DT phase diagram. This means that the two diagrams are essentially different ways of visualizing the same data, only for slightly different ranges of the sparsity parameter. We find the RSS diagram more intuitive to use for our purposes, because the quantities of interest — the relative sparsity and sampling — can be read directly on the coordinate axes.
3. The DT phase diagram is typically used for studying randomly generated matrix instances A in addition to image instances x_{orig}, with the objective to understand recoverability over a whole matrix ensemble. For CT, on the other hand, we are interested in recoverability with a fixed matrix reflecting the given data acquisition geometry of the scanner in question, and the RSS diagram provides a mean for studying this situation: It attempts to answer how many views suffice for recovery of a phantom with a given (relative) sparsity for using a *fixed choice* of system matrix.

For these reasons, the remaining part of the article will solely use the RSS diagram to visualize the recovery results.

4.4. Dependence on image size. To study how the RSS diagram depends on image size, we construct additional diagrams for $N_{\mathrm{side}} = 32$ and 128. For $N_{\mathrm{side}} = 32$ we can use the same relative sampling as for $N_{\mathrm{side}} = 64$ by taking $N_{\mathrm{v}} = 1, 2, \ldots, 16$, since the matrix becomes full-rank at $N_{\mathrm{v}}^{\mathrm{suf}} = 13$. For $N_{\mathrm{side}} = 128$ we have $N_{\mathrm{v}}^{\mathrm{suf}} = 51$, so by taking $N_{\mathrm{v}} = 4, 8, \ldots, 64$ we obtain approximately the same relative sampling.

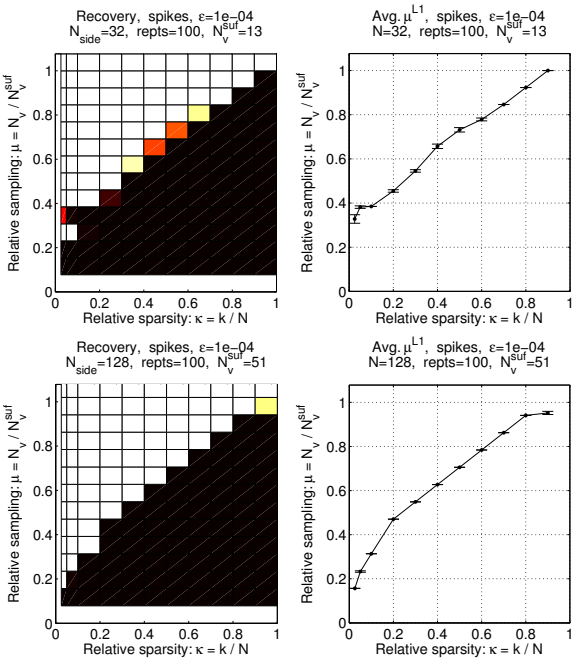


FIG. 4.5. *RSS diagram dependence with image size. Top:* $N_{\text{side}} = 32$. *Bottom:* $N_{\text{side}} = 128$. *Left: recovery of instances. Right: average relative sampling for recovery and* 99% *confidence intervals.*

The two additional RSS diagrams are shown in Figure 4.5. Overall, we see the same monotone increase in $\mu^{\mathbf{L1}}$ with increasing κ. For $N_{\text{side}} = 32$, however, the transition from non-recovery to recovery is slightly more gradual (wider confidence intervals), and the cases with the smallest κ have a larger $\mu^{\mathbf{L1}}$. These differences for are most likely caused by discretization effects.

An interesting phenomenon occurs at $\kappa = 0.025$ and 0.05, where the value $\mu = 0.23$ is sufficient for recovery, but adding one more view to obtain $\mu = 0.31$ destroys recovery. This seemingly counter-intuitive phenomenon is explained by the geometry underlying the data acquisition: Going from 3 to 4 views is not done by including an additional view to the existing views; rather the 4 views are distributed equi-angularly around the image, and hence the two system matrices are entirely different. For this problem, 3 views provide enough data to recover the image, whereas the 4 views are insufficient. The reason is the relatively larger null space of the 4-view matrix; this matrix is 256×812 and has rank 244, whereas the 3-view matrix is 192×812 and has rank 190, i.e., closer to full row-rank.

The RSS diagram for $N_{\text{side}} = 128$ is similar to the $N_{\text{side}} = 64$ case, except for generally sharper transition as well as slightly better recovery at the extreme κ-values. Moreover, $N_{\text{side}} = 64$ is sufficiently large, with the possible exception of the very low values of κ, to give representative results that can be extrapolated to predict the sparsity-sampling relation at larger N_{side}.

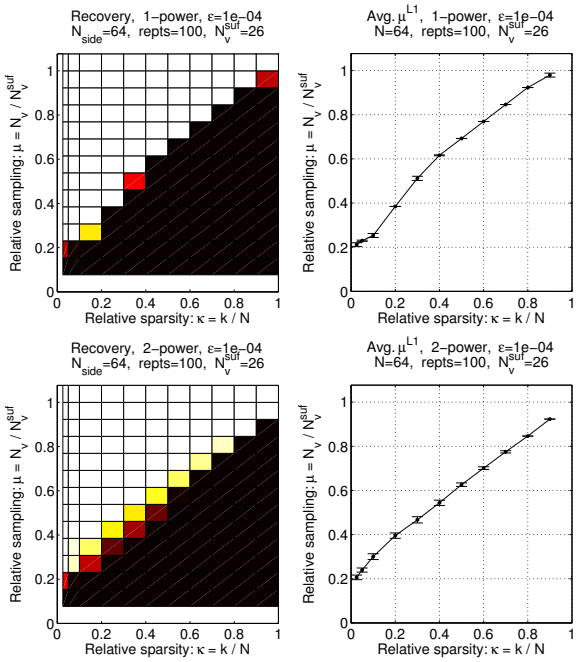


FIG. 4.6. *RSS diagrams for classes* **1-power** *(top) and* **2-power** *(bottom).*

4.5. Dependence on the phantom class. As argued in Section 2.2, we cannot expect recovery of *all* k-sparse images at a given relative sampling — at least only with *very few* non-zeros and unfavorably large relative sampling, due to the existence of pathological phantoms violating otherwise typical sufficient sampling. Hence, we study recoverability only for well-defined classes of phantom images.

Figure 4.6 shows RSS diagrams for the **1-power** and **2-power** classes for $N_{\text{side}} = 64$. Comparing with the **spikes** RSS diagrams in Figure 4.4 we observe similar trends. For **1-power** the transition from non-recovery to recovery occurs at the same (κ, μ)-values, and is almost as sharp as for the **spikes** class. For **2-power** the transition is more gradual, and occurs at lower μ-values for the mid and upper range of κ. We conclude that, on average, a smaller number of views suffices but the in-class recovery variability is larger. Thus, while recoverability is clearly tied to sparsity, the structure of the non-zero pixel locations also plays a role. The RSS diagram can be used to study variation with structure and to determine if two classes have similar recoverability.

To further study how the recoverability depends on image class, we consider the RSS diagrams for the **signedspikes** class in Figure 4.7. Here, the transition is very sharp and occurs at much larger μ-values than for the **spikes** class in Figure 4.4. For example, at $\kappa = 0.4$ **spikes** has average $\mu^{\mathbf{L1}} = 0.62$ compared to 0.77 for **signedspikes**. At $\kappa = 0.8$ **spikes** still has undersampled recovery, although only at average $\mu^{\mathbf{L1}} = 0.92$, compared to no undersampling admitted for **signedspikes**. We conclude that **signedspikes** images are harder to recover, and the RSS diagram allows us to *quantify how much* harder, which we consider a useful property.

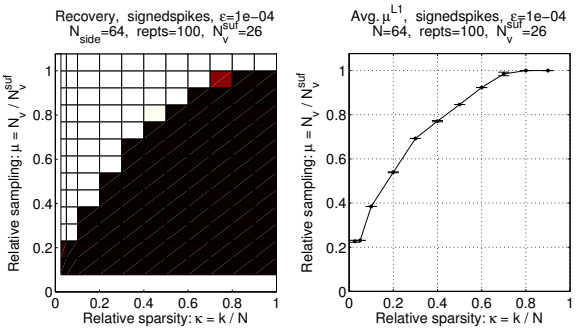


FIG. 4.7. *RSS diagrams for the class* **signedspikes**.

4.6. Robustness to noise. The main focus of the present work is to study the sparsity-sampling relation using the RSS diagram in the ideal noise-free data case. A natural question, however, is whether and how the results generalize in the case of realistic noisy data. We consider therefore the reconstruction problem $\mathbf{L1}_\eta$.

Noise and inconsistencies in CT data are complex subjects arising from many different sources including scatter and preprocessing steps applied to the raw data before the reconstruction step. A comprehensive CT noise model is not our goal as that would necessarily be very application-specific; rather we wish to investigate how the RSS diagram can generalize to non-ideal data. Furthermore, reconstruction from noisy data requires a selection of η and it is well-known that the optimal η is data-dependent.

We model each CT view to have the same fixed x-ray exposure by letting the data in each view b_p, $p = 1, \ldots, N_\mathrm{v}$ be perturbed by the additive zero-mean Gaussian noise vector e_p of constant magnitude $\|e_p\|_2 = \delta$, $p = 1, \ldots, N_\mathrm{v}$. Hence, the noisy data are $b = Ax_\mathrm{orig} + e$, where e is the concatenation of noise vectors for all views. We use three noise levels, $\delta = 10^{-4}, 10^{-2}, 10^{0}$, corresponding to relative noise levels $\|e\|_2/\|Ax_\mathrm{orig}\|_2$ of 1.6%, 0.016% and 0.00016%. We reconstruct using $\mathbf{L1}_\eta$ with $\eta = \|e\|_2 = \sqrt{N_\mathrm{v}} \cdot \delta$ and show the relative reconstruction errors from (3.1) in Figure 4.8.

For $\delta = 10^{-4}$ and 10^{-2} the sudden error drop when the image is recovered is observed at the same number of views as in the noise-free case. The limiting reconstruction error is now governed by the choice of δ and not by the numerical accuracy of the algorithm as in the noise-free case. For the high noise level of $\delta = 10^0$ no sudden error drop can be observed. However, the reconstruction error does continue to decay after the recovery view number seen at the lower noise levels and approach a limiting level consistent with the lower noise-level error curves.

In order to set up RSS diagrams we must choose appropriate thresholds ϵ to match the limiting reconstruction error at each noise level. In the noise-free case we used $\epsilon = 10^{-4}$ chosen to be roughly the midpoint between the initial and limiting errors of the order of 10^0 and 10^{-8}, respectively. Using the same strategy we obtain thresholds $10^{-2.5}$, $10^{-1.5}$, $10^{-0.5}$ for increasing noise level δ. The resulting RSS diagrams are shown in Figure 4.9 (the average-case diagrams have been left out for brevity). The low-noise RSS diagram (to the left) is essentially unchanged from the noise-free case in Figure 4.4. With increasing noise level we see that the location of the transition

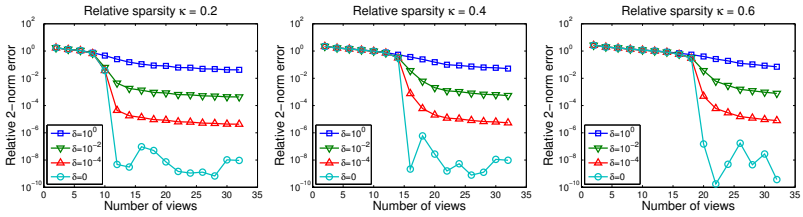


FIG. 4.8. *Numerical recovery measure from Figure* (3.1) *vs. view numbers for* $\mathbf{L1}_\eta$ *reconstruction at different noise levels. Same* **spikes** *phantom instances as in Fig. 4.2 with relative sparsity values* $\kappa = 0.2,\ 0.4$ *and* 0.6. *The "*$\delta = 0$*" case is the noise-free* **L1** *result for reference.*

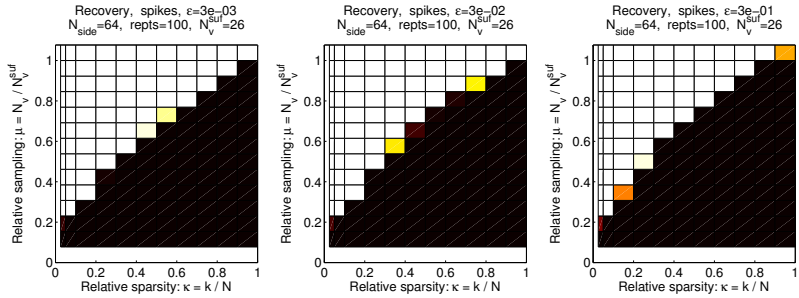


FIG. 4.9. *RSS diagrams for* $\mathbf{L1}_\eta$ *reconstruction of the class* **spikes** *at different noise levels, from left to right,* $\delta = 10^{-4},\ 10^{-2},\ 10^{0}$.

is gradually shifted to higher μ values for the medium and large κ values. At the high noise level and the largest $\kappa = 0.9$ (rightmost plot, upper right corner) we even see that there are instances that are not recovered (to the chosen threshold ϵ) at the sufficient view number N_v^{suf} for having a full-rank system matrix A.

We conclude that sparsity-sampling relation revealed by the RSS diagram in the noise-free case is robust to low levels of Gaussian noise. For medium and high levels of noise, the RSS diagram shows that a sharp transition continues to hold (for the particular noise considered) but the location of the transition changes to require more data for accurate reconstruction.

5. Discussion of our methodology. Our CT simulation studies show that for several phantom classes with different sparsity structure it is possible to observe a sharp transition from non-recovery to recovery in the sense that same-sparsity phantom realizations require essentially the same sampling for accurate reconstruction. In light of the lack of theoretical recovery results mentioned in Section 2.2, we find it surprising that such a sharp transition exists holds for CT matrices for a real sampling configuration without any artificial randomness.

5.1. Limitations and extensions. While the present studies consider a simplified CT system we believe that our results can provide some guidance on the sparsity-sampling relation for a realistic CT system. Our intention here is to take a first step in this direction by proposing to carry out studies of particular systems of interest, and to provide an analysis for a simple but easily generalizable set-up. Our quantita-

tive conclusions are, of course, only valid for the specific geometries, algorithms, data, and phantom classes. Specific applications may call for modifications to our proposed set-up, for instance, the 2-norm metric (3.1) may not be appropriate for evaluating the practical utility of an image in a specific application [2].

Insisting on using a robust optimization algorithm limits the possible image size in the simulations; with MOSEK, we found $N_{\text{side}} > 128$ to be impractical. Therefore, we emphasize that the use of MOSEK in the present paper is to ensure *robust* reconstruction, and we do not advocate MOSEK for larger systems than the ones considered here. Faster algorithms that are applicable to problems of larger scale exist, and potentially we only pay a price of reduced robustness. However, in CT systems of realistic size the number of variables can easily exceed a million in 2D, and be much larger in 3D. Even with the fastest algorithms currently available, a single realistically sized **L1** reconstruction can take hours to days to compute. It will be a daunting task to run the large ensemble reconstructions required for a reliable RSS diagram, and in practice it will likely still be necessary to study a smaller-than-realistic system. It may be more advantageous to stick with the robust algorithm and smaller phantoms for ensuring an accurate solution and consider a larger ensemble in order to further increase the reliability of the RSS diagram. Also, the possibility to extrapolate RSS diagrams to larger image sizes as observed in Section 4.4 reduces the need for studies of larger systems. Since it appears that the relation between relative sparsity and relative sampling for recovery holds across the image size N_{side}, we do not need to study larger N_{side} but can simply extrapolate the relation to more realistic values of N_{side}.

5.2. Future work. The RSS diagram allows for generalization to increasingly realistic set-ups. For example, other phantom classes can be considered with sparsity in other representations, such as in the gradient, and the penalty function can be changed to enforce the expected kind of sparsity. Other kinds of noise and inconsistencies can be introduced in the data and the system can be changed, e.g., to a limited angle CT problem. Making such generalizations might require modifications of the sparsity measure and the recovery criterion.

Our earlier studies of TV reconstructions [17] seem to show a relation between gradient image sparsity and sufficient sampling, but due to the complexity of the test problems we found it difficult to establish any quantitative relation. An investigation based on the RSS diagram could provide more structured insight. For instance, we might learn that TV reconstructions of a class of "blocky" phantoms exhibit a well-defined recovery-curve similar to the ones in the present study.

We always face the problem of possible non-unique solutions to **L1**, leading to RSS and DT phase diagrams that, in principle, depend on the particular choice of optimization algorithm. We expect that **L1**-uniqueness can be studied by numerically verifying a set of necessary and sufficient conditions [14]. We did not pursue that idea in the present work in order to focus on an empirical approach easily generalizable to other penalties, such as TV, for which similar conditions may not be available.

Another interesting future direction is to study the in-class recovery variability, i.e., why the **2-power** class transition from non-recovery to recovery is more gradual. Would it be possible to identify differences between instances that were recovered and ones that were not, e.g., in the spatial location of the non-zero pixels, or in the histograms of pixel values? This could lead to subdividing the phantom class into partitions each having sharper transitions occurring at different relative sampling values and thereby an even better understanding of what factors influence the recoverability.

6. Conclusion. Inspired by the Donoho-Tanner phase diagram we devised a relative sparsity-sampling (RSS) diagram for empirically studies of recoverability in sparsity-exploiting x-ray CT image reconstruction. We focused on pixel sparsity and 1-norm-based reconstruction, but our approach is not limited to sparsity in a specific domain or reconstruction by solving a specific optimization problem.

Our numerical simulations using the RSS diagram demonstrate a pronounced relation between image sparsity and the number of projections needed for recovery, for a range of image classes without and with structure and classes with signed and unsigned pixel values. In the majority of the studied cases, we found a sharp transition from non-recovery to recovery — a result that hitherto, to our knowledge, has not been established for CT. The sharp transition allows for quantitatively predicting the number of projections that, on average, suffices for **L1**-recovery of phantom images from a specific class, or conversely, to determine the maximal sparsity of an image that can be recovered for a certain number of views. We saw that the transition from non-recovery to recovery is independent of the image size and robust to small amounts of additive Gaussian noise. With these initial results we have taken a step towards better quantitative understanding of the recoverability from undersampled measurements in x-ray CT, and additionally we provide a tool for determining similar answers for increasingly realistic systems.

In summary, we believe that the RSS diagram can provide quantitative insight into sparsity-exploiting reconstruction because 1) it provides a structured framework for establishing and quantifying the relation between sparsity and sufficient sampling for a particular system, 2) it does not rely on existence of theoretical results guaranteeing solution uniqueness, and 3) it allows the study of realistically-sized systems through extrapolation from smaller systems for which reconstruction of ensembles is feasible.

Acknowledgment. We thank the associate editor and two anonymous referees for providing valuable comments that significantly improved the article. This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. JSJ also acknowledges support from The Danish Ministry of Science, Innovation and Higher Education's Elite Research Scholarship. This work was supported in part by NIH R01 grants CA158446, CA120540 and EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, *Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery*, Appl. Comput. Harmon. Anal., 26 (2009), pp. 283–290.

[2] H. H. Barrett and K. J. Myers, *Foundations of Image Science*, John Wiley & Sons, Hoboken, NJ, 2004.

[3] M. Beister, D. Kolditz, and W. A. Kalender, *Iterative reconstruction methods in X-ray CT*, Physica Med., 28 (2012), pp. 94–108.

[4] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, *Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT*, Phys. Med. Biol., 55 (2010), pp. 6575–6599.

[5] A. M. Bruckstein, D. L. Donoho, and M. Elad, *From sparse solutions of systems of equations to sparse modeling of signals and images*, SIAM Rev., 51 (2009), pp. 34–81.

[6] E. J. Candès, J. Romberg, and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory, 52 (2006), pp. 489–509.

[7] E. J. Candès and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory, 51 (2005), pp. 4203–4215.

[8] D. Donoho and J. Tanner, *Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing.*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 367 (2009), pp. 4273–4293.

[9] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory, 52 (2006), pp. 1289–1306.

[10] D. L. Donoho and M. Elad, *Optimally sparse representation in general (non-orthogonal) dictionaries via L1 minimization*, Proc. Natl. Acad. Sci. USA, 100 (2003), pp. 2197–2202.

[11] C. Dossal, G. Peyré, and J. Fadili, *A numerical exploration of compressed sampling recovery*, Linear Algebra Appl., 432 (2010), pp. 1663–1679.

[12] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*, Springer, New York, NY, 2010.

[13] L. A. Feldkamp, L. C. Davis, and J. W. Kress, *Practical cone-beam algorithm*, J. Opt. Soc. Amer. A, 1 (1984), pp. 612–619.

[14] M. Grasmair, M. Haltmeier, and O. Scherzer, *Necessary and sufficient conditions for linear convergence of L1-regularization*, Comm. Pure Appl. Math., 64 (2011), pp. 161–182.

[15] X. Han, J. Bian, D. R. Eaker, T. L. Kline, E. Y. Sidky, E. L. Ritman, and X. Pan, *Algorithm-enabled low-dose micro-CT imaging*, IEEE Trans. Med. Imaging, 30 (2011), pp. 606–620.

[16] P. C. Hansen and M. Saxild-Hansen, *AIR Tools – A MATLAB package of algebraic iterative reconstruction methods*, J. Comput. Appl. Math., 236 (2012), pp. 2167–2178.

[17] J. S. Jørgensen, E. Y. Sidky, and X. Pan, *Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in x-ray CT*, IEEE Trans. Med. Imaging, 32 (2013), pp. 460–473.

[18] M. Li, H. Yang, and H. Kudo, *An accurate iterative reconstruction algorithm for sparse objects: application to 3D blood vessel reconstruction from a limited number of projections*, Phys. Med. Biol., 47 (2002), pp. 2599–2609.

[19] MOSEK ApS, *MOSEK Optimization Software (www.mosek.com)*.

[20] F. Natterer, *The Mathematics of Computerized Tomography*, John Wiley & Sons, New York, NY, 1986.

[21] X. Pan, E. Y. Sidky, and M. Vannier, *Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?*, Inverse Problems, 25 (2009), p. 123009.

[22] S. Petra and C. Schnörr, *TomoPIV meets compressed sensing*, Pure Math. Appl. (PU.M.A.), 20 (2009), pp. 1737–1739.

[23] M. E. Pfetsch and A. M. Tillmann, *The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing*, Arxiv preprint arXiv:1205.2081, (2012).

[24] N. Pustelnik, C. Dossal, F. Turcu, Y. Berthoumieu, and P. Ricoux, *A greedy algorithm to extract sparsity degree for L1/L0-equivalence in a deterministic context*, in Proceedings of EUSIPCO, Bucharest, Romania, 2012.

[25] I. Reiser and R. M. Nishikawa, *Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise*, Med. Phys., 37 (2010), pp. 1591–1600.

[26] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelriess, *Improved total variation-based CT image reconstruction applied to clinical data*, Phys. Med. Biol., 56 (2011), pp. 1545–1561.

[27] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Phys. D, 60 (1992), pp. 259–268.

[28] E. Y. Sidky, C.-M. Kao, and X. Pan, *Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT*, J. X-Ray Sci. Technol., 14 (2006), pp. 119–139.

[29] E. Y. Sidky and X. Pan, *Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization*, Phys. Med. Biol., 53 (2008), pp. 4777–4807.

[30] L. Yu, X. Liu, S. Leng, J. M. Kofler, J. C. Ramirez-Giraldo, M. Qu, J. Christner, J. G. Fletcher, and C. H. McCollough, *Radiation dose reduction in computed tomography: techniques and future perspective.*, Imaging Med., 1 (2009), pp. 65–84.

Appendix D

Few-view single photon emission computed tomography (SPECT) reconstruction based on a blurred piecewise constant object model

---

Submitted to Phys. Med. Biol., 2012.

P. A. Wolf, **J. H. Jørgensen**, T. G. Schmidt and E. Y. Sidky

Few-view single photon emission computed tomography (SPECT) reconstruction based on a blurred piecewise constant object model

Paul A Wolf[1] , Jakob H Jørgensen[2], Taly G Schmidt[1] and Emil Y Sidky[3]

[1]Department of Biomedical Engineering, Marquette University, 1515 W. Wisconsin Ave. Milwaukee, WI 53233, USA

[2]Department of Informatics and Mathematical Modeling, Technical University of Denmark, Richard Petersens Plads, Building 321, 2800 Kgs. Lyngby, Denmark

[3]Department of Radiology, University of Chicago, 5841 S. Maryland Ave., Chicago, IL 60637, USA

Email: paul.wolf@marquette.edu, jakj@imm.dtu.dk, tal.gilat-schmidt@marquette.edu and sidky@uchicago.edu

Short title. Few-view SPECT reconstruction based on a blurred piecewise constant object model

Abstract. A sparsity-exploiting algorithm intended for few-view Single Photon Emission Computed Tomography (SPECT) reconstruction is proposed and characterized. The algorithm models the object as piecewise constant subject to a blurring operation. To validate that the algorithm closely approximates the true object in the noiseless case, projection data were generated from an object assuming this model and using the system matrix. Monte Carlo simulations were performed to provide more realistic data of a phantom with varying smoothness across the field of view. Reconstructions were performed across a sweep of two primary design parameters. The results demonstrate that the algorithm recovers the object in a noiseless simulation case. While the algorithm assumes a specific blurring model, the results suggest that the algorithm may provide high reconstruction accuracy even when the object does not match the assumed blurring model. Generally, increased values of the blurring parameter and TV weighting parameters reduced noise and streaking artifacts, while decreasing spatial resolution. As the number of views decreased from 60 to 9 the accuracy of images reconstructed using the proposed algorithm varied by less than 3%. Overall, the results demonstrate preliminary feasibility of a sparsity-exploiting reconstruction algorithm which may be beneficial for few-view SPECT.

Keywords: sparsity-exploiting reconstruction, SPECT, tomographic reconstruction

PACS classification numbers: 87.57.nf, 87.57.uh

1. Introduction

Single Photon Emission Computed Tomography (SPECT) provides noninvasive images of the distribution of radiotracer molecules. Dynamic Single Photon Emission Computed Tomography provides information about tracer uptake and washout from a series of time-sequence images. Dynamic SPECT acquisition methods measuring time activity curves on the order of minutes have been developed (Gullberg *et al* 2010, Gullberg 2004). However, the dynamic wash-in wash-out of some tracers occurs over a period of just several seconds, requiring better temporal sampling. Stationary ring-like multi-camera systems are being developed to provide rapid dynamic acquisitions with high temporal sampling (Beekman *et al* 2005, Furenlid *et al* 2004, Beekman and Vastenhouw 2004). Reducing the number of cameras reduces the cost of such systems but also reduces the number of views acquired, limiting the angular sampling of the system. Novel few-view image reconstruction methods may be beneficial and are being investigated for the application of dynamic SPECT (Ma *et al* 2012).

The feasibility of reconstructing from angularly undersampled, or few-view data, has recently been explored for CT (Sidky and Pan 2008, Chen et al 2008, Duan et al 2009, Ritschl et al 2011, Sidky et al 2006). These investigations are based on exploitation of gradient-magnitude sparsity, an idea promoted and theoretically investigated in the field of Compressed Sensing (CS). Few-view, sparsity-exploiting CT reconstruction algorithms promote gradient-magnitude sparsity by minimizing image total variation (TV). Success of these algorithms in allowing sampling reduction follows from an object model which is approximately piecewise constant, a model that may not apply well for SPECT objects. The SPECT object function quantifies the physiological uptake of a radiolabelled tracer in the body. In some applications, the transition between different uptake regions in the SPECT object is expected to be smoother than the transition between X-ray attenuation coefficients in the CT object. The goal of this work is to modify the idea of exploiting gradient-magnitude sparsity to allow for smoother transitions between regions of approximately constant values of tracer concentration.

This paper proposes an iterative algorithm for few-view SPECT reconstruction that allows for smoothed step-like variation within the object by phenomenologically modeling the SPECT object as a blurred version of a

piecewise constant object. Using this model, a first-order primal-dual technique is implemented as an iterative procedure (Chambolle and Pock 2011, Sidky *et al* 2012). The purpose of this study was to characterize the performance of the algorithm under varying sampling and noise conditions, including cases where the object does not match the phenomenological model. Images reconstructed by Maximum-Likelihood Expectation Maximization (MLEM) serve as a reference. The article is organized as follows: Section 2 provides the image reconstruction theory and algorithm. Sections 3 and 4 demonstrate the algorithm with data generated using the system matrix and with data generated from a realistic Monte Carlo simulation of a SPECT system, respectively. Section 5 summarizes the results.

2. The algorithm

The iterative image reconstruction algorithm (IIR) is designed by defining an optimization problem which implicitly specifies the object function based on a realistic data model and a model for object sparsity. In this preliminary investigation, the specified optimization problem is solved in order to characterize its solution and the solution's appropriateness for few-view/dynamic SPECT imaging. Future work will consider algorithm efficiency by designing IIR for approximate solution of the proposed optimization problem.

2.1. The SPECT optimization problem

The proposed SPECT optimization problem is formulated as an unconstrained minimization of an objective function which is the sum of a data fidelity term and an image regularity penalty. The design of both terms expresses the proper SPECT noise model and a modified version of gradient-magnitude object sparsity. We first describe how standard gradient-magnitude sparsity is incorporated into a SPECT optimization problem, and then we present our modified optimization which accounts for the smoother variations expected in a SPECT object function.

2.1.1. Unconstrained minimization for gradient-magnitude sparsity exploiting SPECT IIR. In expressing the SPECT data fidelity term, the data are modeled as a Poisson process the mean of which is described by the following linear system of equations:

$$\boldsymbol{g} = \boldsymbol{H}\boldsymbol{f} \tag{1}$$

where $\boldsymbol{H}$ is the system matrix that describes the probability that a photon emitted from a certain location in the object vector, $\boldsymbol{f}$, contributes to the measured data vector, $\boldsymbol{g}$, at a certain location. Iterative tomographic image reconstruction techniques such as MLEM and Ordered Subset Expectation Maximization (OSEM) maximize the log-likelihood of this Poisson random variable (Shepp and Vardi 1982, Hudson and Larkin 1994, Vandenberghe *et al* 2001). This is equivalent to minimizing the Kullback-Leibler (KL) data divergence (D_{KL}) (Barrett and Myers 2004). For the present application of few-view SPECT, the data are acquired over too few views to provide a unique maximum likelihood image. In the limit of infinite photon counts and assuming that the mean model in (1) perfectly describes the imaging system, the underlying object function still cannot be determined because (1) is underdetermined.

In order to arrive at a reasonable solution, additional information or assumptions on the object function are needed. Recently, exploitation of gradient-magnitude sparsity has received much attention and has been implemented in IIR for few-view CT (Chen *et al* 2008, Sidky *et al* 2009). This idea is an example of a general strategy under much recent investigation in CS, where sampling conditions are based on some form of identified sparsity in the image. In our application the strategy calls for narrowing the solution space to only images that exactly solve our linear model in (1). Among those images, the solution with the lowest TV is sought. In practice, this solution can be obtained approximately by combining a data fidelity term with a TV penalty, where the combination coefficient in front of the TV penalty is vanishingly small. The TV-D_{KL} sum yields the following minimization:

$$\mathit{minimize}_f \ \{D_{\mathrm{KL}}(\boldsymbol{g}, \boldsymbol{H}\boldsymbol{f}) + \gamma \left\| (|\boldsymbol{D}\boldsymbol{f}|) \right\|_1 \}, \tag{2}$$

where $\boldsymbol{D}$ is a discrete gradient operator and γ is a weighting parameter. For sparsity-exploiting IIR, γ is chosen so that the data fidelity term far outweighs the TV-term. The role of the TV term is simply to break the degeneracy in the objective function among all solutions of (1).

The success of TV minimization for few-view CT IIR relies on the assumption that the X-ray attenuation coefficient map is approximately piecewise constant. Directly promoting sparsity of the gradient-magnitude image may not be as beneficial for SPECT, as in some cases tracer uptake may vary smoothly within objects, and borders of objects may show a smoothed step-like dependence. For example, some regions of the heart are supplied by a single coronary artery while other regions are supplied by multiple coronary arteries (Donato *et al* 2012, Pereztol-Valdés *et al* 2005). Thus, cardiac perfusion studies may be one application for which the blurred piecewise constant model is appropriate. As another example, tumor vascularization is heterogeneous, with vascularization often varying from the tumor center to the periphery (Jain 1988). Therefore, our goal here is to find a sparsity-exploiting formulation which allows some degree of smoothness between regions with different uptake.

2.1.2. Unconstrained minimization for sparsity exploiting IIR using a blurred piecewise constant object model. In this work the TV minimization detailed in (2) is modified to allow for rapid but smooth variation by phenomenologically modeling objects as piecewise constant subject to a shift-invariant blurring operation.

The additional blurring operation can be incorporated into the framework developed above by minimizing the weighted sum of the TV of an intermediate piecewise constant object estimate and D_{KL} between the measured data and the projection data of the blurred object estimate. The modified TV-minimization problem becomes:

$$\mathit{minimize}_f \ \{D_{\mathrm{KL}}(\boldsymbol{g}, \boldsymbol{Hu}) \ + \gamma \left\| (|\boldsymbol{Df}|) \right\|_1 \}, \tag{3}$$

where $\boldsymbol{u}$ is the object estimate and $\boldsymbol{f}$ is an intermediate image with sparse gradient-magnitude. These are related by $\boldsymbol{u} = \boldsymbol{MGMf}$, where $\boldsymbol{M}$ is a support preserving image mask and $\boldsymbol{G}$ is a Gaussian blurring operation with standard deviation r. The operators $\boldsymbol{M}$ and $\boldsymbol{G}$ are symmetric so $\boldsymbol{M}^{\mathrm{T}} = \boldsymbol{M}$ and GT=$\boldsymbol{G}$. The operator $\boldsymbol{G}$ extends data outside the physical support of the system assumed by $\boldsymbol{H}$ so the image mask $\boldsymbol{M}$ must be applied before and after

$\boldsymbol{G}$. This optimization problem has two design parameters, γ, which is the weighting of the TV term, and r, which is the standard deviation of the Gaussian blurring kernel. The blurring parameter, r, represents smoothness in the underlying object, as opposed to blurring introduced by the imaging system. When $r = 0$, this formulation defaults to TV minimization problem in (2). If $\gamma = 0$, the formulation described by (3) minimizes D_{KL}, which is implicitly minimized in MLEM. The final image estimate is $\boldsymbol{u}$, the result of blurring and masking the intermediate piecewise constant object, $\boldsymbol{f}$. Minimizing (3) jointly enforces sparsity (by requiring a low TV of $\boldsymbol{f}$) and encourages data match (by requiring a low D_{KL}).

2.2. Optimization algorithm

Only recently have algorithms been developed that can be applied to large-scale, non-smooth convex optimization problems such as that posed by (3). Sidky *et al* (2012) adapts the Chambolle-Pock (CP) algorithm to solve the TV- D_{KL} sum described by (2) (Chambolle and Pock 2011). Applying the model as described above, this prototype can be modified to solve the optimization posed by (3). Pseudo-code describing this algorithm is written below.

Listing 1: Pseudocode of the proposed algorithm

$L := \|(\boldsymbol{HMGM},\boldsymbol{D})\|_2$; $\tau = \sigma = 0.9/L$; $\theta = 1$; $n = 0$

$\boldsymbol{f}_0 := \boldsymbol{f'}_0 := \boldsymbol{p_0} := \boldsymbol{q_0} := \boldsymbol{0}$

Repeat

$\boldsymbol{p}_{n+1} := 0.5(1 + \boldsymbol{p}_n + \sigma\, \boldsymbol{HMGMf'}_n - ((p_n + \sigma \boldsymbol{HMGMf'}_n - \boldsymbol{1})^2 + 4\sigma \boldsymbol{g})^{1/2}$

$\boldsymbol{q}_{n+1} := \gamma(\boldsymbol{q}_n + \sigma \boldsymbol{Df'}_n) / \max(\gamma, |\boldsymbol{q}_n + \sigma \boldsymbol{Df'}_n|)$

$\boldsymbol{f}_{n+1} := \boldsymbol{f}_n - \tau \boldsymbol{MGMH}^T \boldsymbol{p}_{n+1} + \tau div(\boldsymbol{q}_{n+1})$

$\boldsymbol{f'}_{n+1} := \boldsymbol{f}_{n+1} + \theta(\boldsymbol{f}_{n+1} - \boldsymbol{f}_n)$

$n = n+1$

Until *stopping criterion*

This algorithm is a modification of Algorithm 5 described in previous work by Sidky *et al* (2012). The convergence criterion described in that work was used here.

Simulation studies were conducted to characterize the performance of the proposed reconstruction technique over a range of angular sampling conditions, including cases in which the object does not match the phenomenological blurred piecewise constant model. The first simulation study used noiseless data generated from the system forward model to validate that the reconstruction technique closely approximates the true object when the correct blurring and system models are used, and to investigate the effects of the design parameters r and γ. Another study reconstructed data generated by Monte Carlo simulation for a range of sampling and noise conditions and for varying values of algorithm parameters r and γ.

3. Inverse crime simulation study

This study was designed to validate that the reconstruction technique approximates the true object when both the object model and system model are known exactly. The simulated object was generated from the object model and data were generated from the system forward model. Cases such as this, in which the data were produced directly from the model are referred to as the "inverse crime" (Kaipio and Somersalo 2005). This is investigated in the many-view (128 views) and few-view (9 views) cases. We also examine the effects of different blurring models on the gradient-magnitude sparsity of the intermediate object $\boldsymbol{f}$. The algorithm could enable further reductions in sampling if the blurring model increases the gradient-magnitude sparsity compared to the conventional TV minimization term. In order to investigate the performance of the reconstruction with inconsistent data, Poisson noise was added to the data and the study was repeated. We refer to this as the "noisy" case.

3.1. Methods

3.1.1. Phantom. The intermediate piecewise constant object, $\boldsymbol{f}_{\text{true}}$, was defined on a 128 x 128 grid of 1-mm x 1-mm pixels, representing a 6-mm diameter disk embedded in a 76-mm diameter disk. The intensity of the small disk was 2000 arbitrary units and the intensity of the large disk was 200 arbitrary units. A Gaussian blurring

kernel with standard deviation, r_{true} = 0.75 pixels was applied to this intermediate object to generate the ground-truth object. The intermediate object and the output of the blurring operation, $\boldsymbol{u}_{true}$ are shown in figure 1.

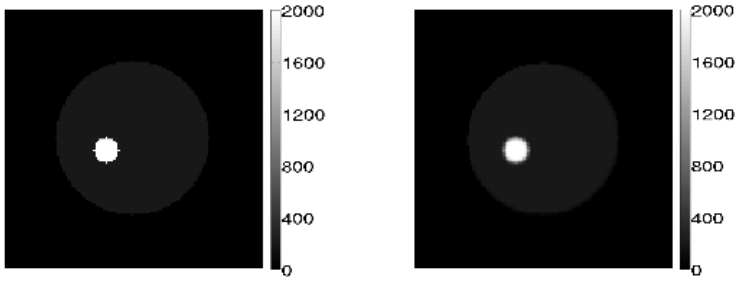


Figure 1. Piecewise Constant Object (left) and Phantom (right) used for simulations that generated data from the system matrix.

3.1.2. Simulation. Projection data of the pixelized ground-truth object was generated from the system matrix. The system matrix was estimated using Siddon's raytracing algorithm for a single-pinhole SPECT system with 3 mm pinhole diameter, 1.0 mm system FWHM, 35 mm pinhole-to-object distance and 63.5 mm pinhole-to-detector distance (Siddon 1985). Projection data were generated and reconstructed using 128 views, 60 views, 21 views, 15 views and 9 views, uniformly distributed around 360 degrees.

A parametric sweep was performed to investigate the effects of the two parameters on the reconstructions: the TV weighting parameter, γ, and the standard deviation of the Gaussian blurring kernel, r. Reconstructions were performed with γ varying from 0.0001 to 1.0 and r varying from 0 to 2.0 pixels. For this case, r_{true} is known to be equal to 0.75. In practice, the amount of smoothness within the underlying object is unknown and may vary across the FOV. In this study, images are reconstructed using a range of r values to quantify the performance of the reconstruction technique for the expected case where the assumed r differs from r_{true}. To reduce the necessary sampling for accurate image reconstruction, a sparse representation of an image must exist. Our proposed reconstruction approach assumes that the gradient-magnitude of the intermediate image $\boldsymbol{f}$ has very few meaningful coefficients. However, using an incorrect blurring model in the reconstruction

may negatively affect the sparsity of the intermediate object, $\boldsymbol{f}$, limiting the benefits of the algorithm. To investigate the effect of the assumed blurring model on the sparsity of the reconstructed intermediate object, $\boldsymbol{f}$, images were reconstructed from 9 and 128 views using a range of r values and sparsity evaluated as the number of coefficients greater than 10% of the maximum coefficient in the gradient-magnitude image of $\boldsymbol{f}$.

To investigate the performance of the reconstruction technique in the presence of noise, simulations varying the number of views and parameter values were repeated with Poisson noise added to the projections generated from the system model. All simulations modeled approximately 1052000 counts, thus the peak number of counts in the 128 view projections was 298 while the peak number of counts in the 9 view projections was 3758. The noisy projection data were also reconstructed with MLEM in order to provide a reference reconstruction for comparison. As will be described in the next section, the correlation coefficient (CC) of the reconstructed image with the true object is used as a metric of accuracy throughout this work. In order to select a comparable stopping iteration for MLEM reconstruction, the CC was calculated at each MLEM iteration and the final image selected as that with the highest CC value.

3.1.3. Metrics. Evaluating the accuracy of the reconstructed object requires a measure of similarity or error between the reconstructed object and the true object. In SPECT imaging, including the Geant4 Application for Tomographic Emission (GATE) simulations proposed in section 4, the reconstructed activity is a scaled version of the true activity, with the scaling factor dependent on the geometric efficiency of the system (Jan *et al* 2004). Our reconstruction methods correct for the spatially varying sensitivity of the SPECT system, as will be described in section 4.1.2. However, a global scaling correction factor is not applied because absolute quantification in SPECT is challenging and may confound the characterization of the algorithm. Therefore, our accuracy metric must provide a meaningful measure of similarity in cases where the scaling factor between the reconstructed and true object is unknown. In this work, reconstruction accuracy was quantified using the correlation coefficient (CC) of the reconstructed image estimate with the true object. CC is defined as

$$CC = \frac{\sum_{k=1}^{M}(u(k)-\bar{u})(u_{\text{true}}(k)-\overline{u_{\text{true}}})}{\{\sum_{k=1}^{M}(u(k)-\bar{u})^2\sum_{k=1}^{M}(u_{\text{true}}(k)-\overline{u_{\text{true}}})^2\}^{1/2}} \quad (4)$$

where u_{true} is the true object, M is the number of voxels and $u(k)$ is the reconstructed object value at voxel k. This metric is commonly used in image registration and is the optimum similarity measure for images that vary by a linear factor (Hill *et al* 2001). This metric allows the quantification of the accuracy of the spatial distribution of the object, without requiring absolute quantitative accuracy. CC is equal to one when the reconstructed object matches the true object. We also quantified the change in CC over the range of studied parameters (r and γ), in order to quantify the sensitivity of the algorithm to parameter selections and to understand the performance of the algorithm when the assumed blurring parameter does not match the true object blur. Spatial resolution in the reconstructed images was quantified as the full-width at 10% of maximum (FW10M) of the central profile through the smaller disk. This measure was used instead of the more common full-width at half maximum (FWHM) because analysis of preliminary reconstructed images indicated that the FWHM was often accurate, even though the extent of the reconstructed object was greater than the true object. The FW10M more accurately quantified this blurring effect. The true object had a FW10M of 12 pixels. Signal-to-noise ratio (SNR) was calculated as the mean of a 3 pixel radius region in the background divided by the standard deviation of the same region.

3.2. Results

3.2.1. Without Poisson noise. We present results of the noiseless case in which the object was constructed from the object model and the projections were determined from the system matrix. The purpose of this study was to confirm that the reconstruction algorithm closely approximated the true object in the noiseless inverse crime case and to examine the effects of the design parameters r and γ as the number of views decreased. Both design parameters were varied and the number of angular samples reduced from 128 views to 9 views.

Reconstructions from 128 angular positions are shown in figure 2, with profiles of selected reconstructions shown in figure 3. Figure 4(a) presents plots of CC over the range of studied γ and r values Reconstruction accuracy (CC) is high (CC > 0.980) for reconstruction using $\gamma < 1.0$, with CC varying by less than 2% for all r investigated. Using $\gamma = 0.0001$ or $\gamma = 0.01$ and using $r = r_{true} = 0.75$, the object is recovered nearly exactly with CC exceeding 0.999 in each case. The FW10M value of the true object is 12 pixels, which is correctly depicted by reconstructions using $r < 1.0$ and $\gamma < 0.1$. The profiles demonstrate decreased amplitude and increased object extent when $\gamma = 0.1$ or $\gamma = 1.0$, suggesting blurring of the object. For the lower γ cases, ring artifacts are visible when r is greater than 1.0.

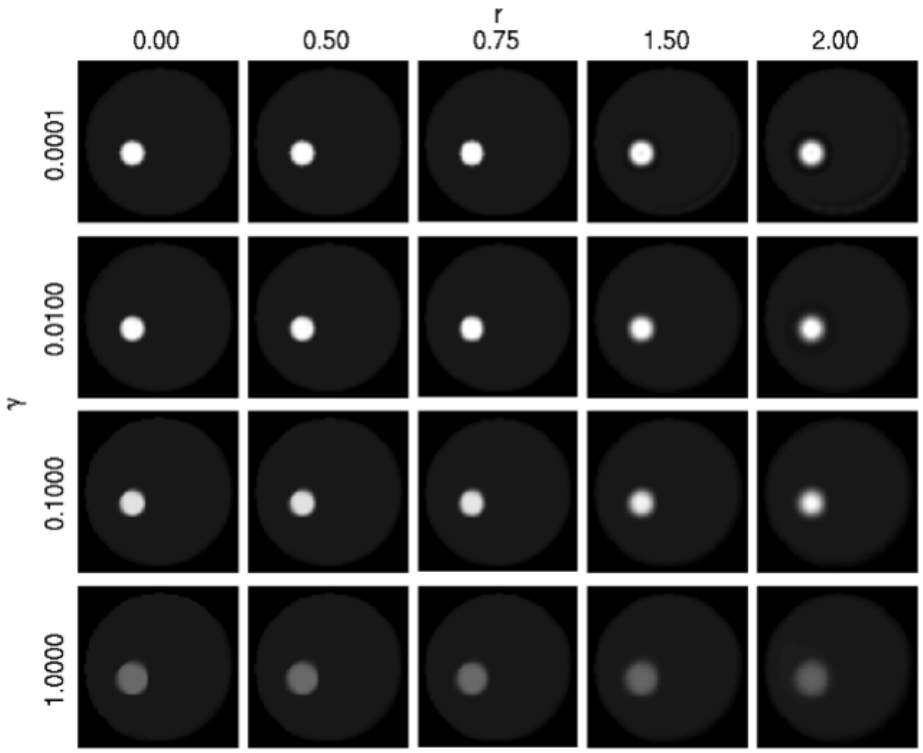


Figure 2. Images reconstructed from 128 views of noiseless inverse crime data using the proposed algorithm with varying values of r and γ.

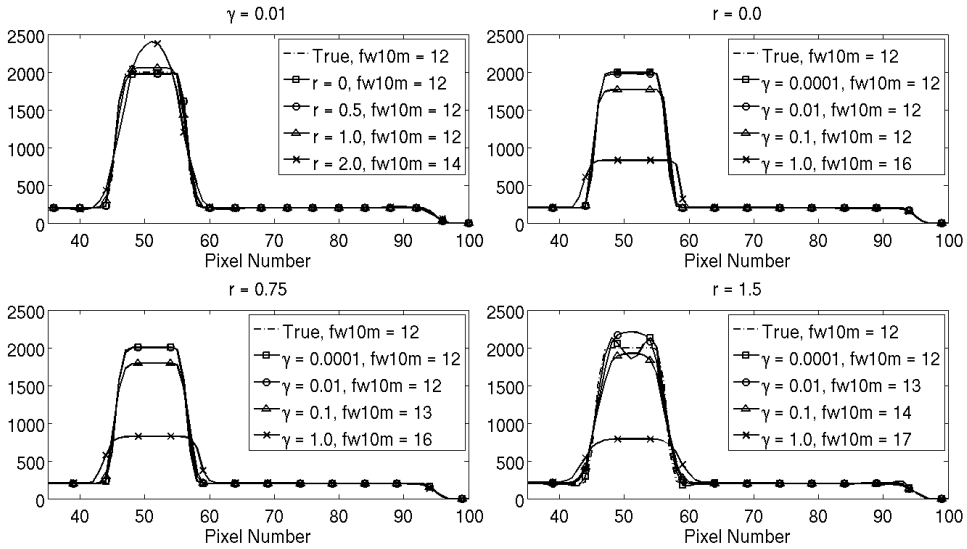


Figure 3. Central diagonal profiles through images reconstructed from noiseless inverse crime data from128 views using the proposed algorithm with varying values of r and γ.

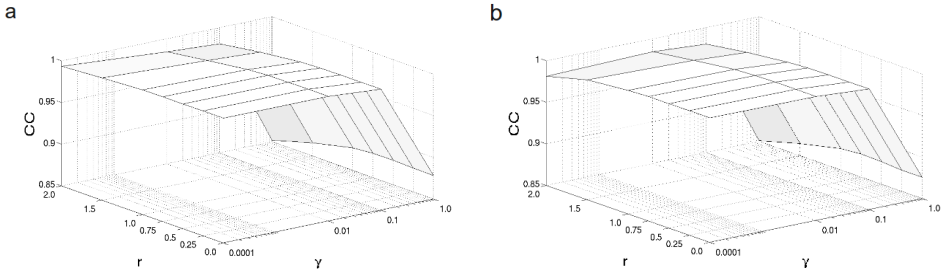


Figure 4. Plots depicting the CC over the range of studied r and γ parameters of images reconstructed from noiseless inverse crime data from 128 views (a) and 9 views (b).

The few-view case demonstrated similar trends, as shown in figures 4(b), 5 and 6. The object is nearly exactly recovered when $\gamma = 0.01$ and $r = 0.75$ is used. Using $\gamma = 0.01$, the CC varied by 1.5% (CC ranging from 0.984 - 0.999) across the range of studied r values. Over the parameter set studied, CC varied between 0.878 and 0.999 depending on the value of γ used in reconstruction, with higher γ values resulting in lower CC. In addition

to lower CC, images reconstructed with $\gamma = 1.0$ demonstrated reduced contrast and increased FW10M results, suggesting increased blurring. The FW10M value of the true object was 12, which was depicted in all reconstructions with $r \leq 1.0$ and $\gamma < 0.1$. As r increased beyond 1.0, the peak value increased and the profiles demonstrate larger spread, resulting in the lowest CC for reconstructions with $r = 2.0$. Overall, in both the 128 and 9-view case, CC values demonstrated a larger range over the set of γ values compared to r values, suggesting that the reconstruction technique is more sensitive to the selection of γ than r.

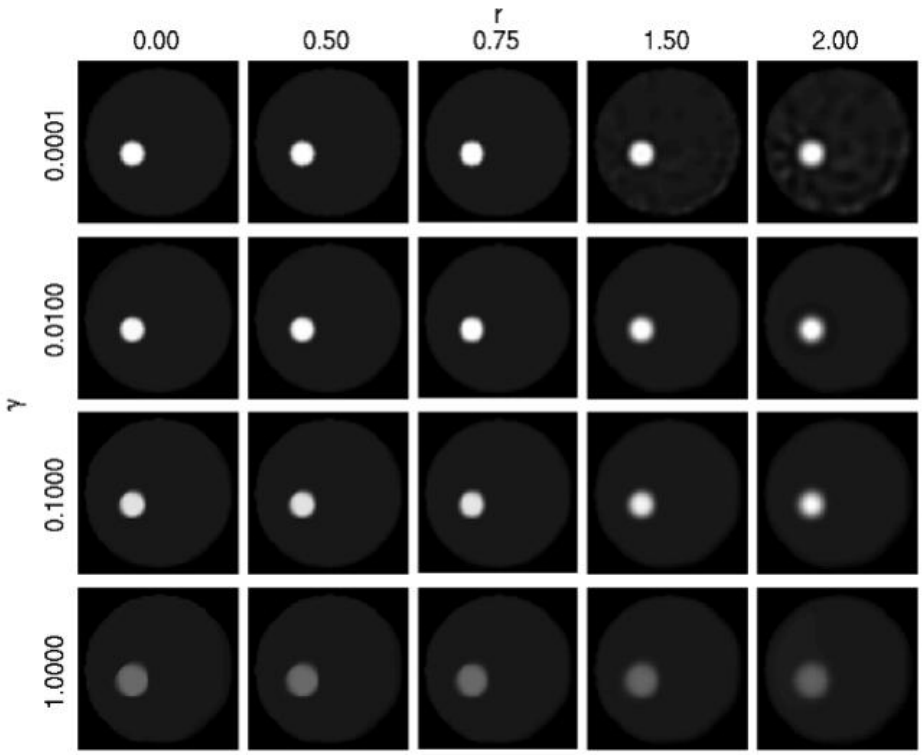


Figure 5. Images reconstructed from 9 views of noiseless inverse crime data using the proposed algorithm with varying values of r and γ.

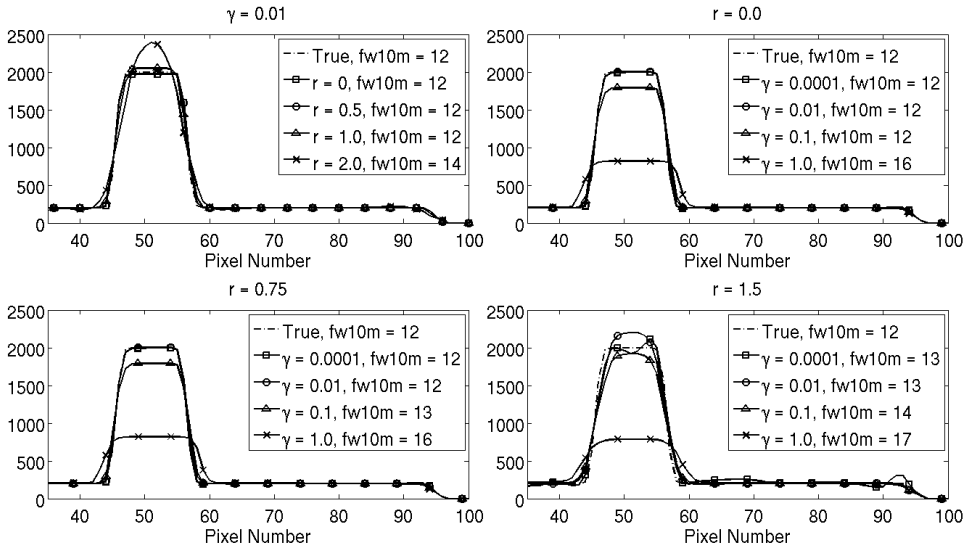


Figure 6. Central diagonal profiles through images reconstructed from noiseless inverse crime data from 9 views using the proposed algorithm with varying values of r and γ.

Evaluating gradient magnitude sparsity of the intermediate image. This section evaluates the sparsity of the intermediate image $\boldsymbol{f}$ reconstructed from many-view and few-view data. Figure 7 shows images of the intermediate image, $\boldsymbol{f}$, reconstructed from both 128 views and 9 views using different r and $\gamma = 0.0001$. Each image is captioned by its sparsity value (number of meaningful coefficients).

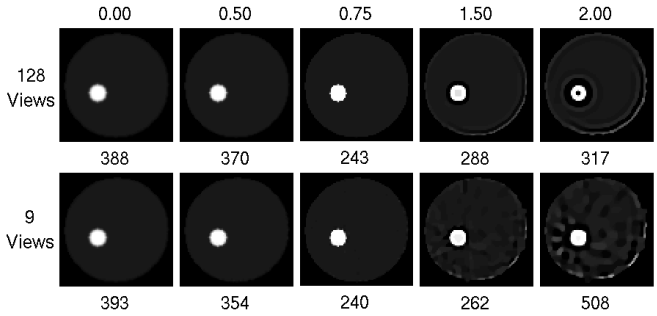


Figure 7. Intermediate images $\boldsymbol{f}$ and the number of meaningful sparsity coefficients reconstructed from128 and 9 noiseless inverse crime data using $\gamma = 0.0001$.

In both the many-view and few-view case, the image reconstructed with the true blurring model (r = 0.75) was the most sparse and as the r assumed by the algorithm diverged from r_{true} the images became less sparse. This indicates that using the correct blurring model may allow the greatest sampling reductions. Additionally, underestimating r leads to a gradual increase in the number of meaningful coefficients. In the few-view case, over-estimating r leads to a rapid increase in the number of meaningful coefficients, reflected by the fact that new structure enters the image. These artifacts survive the blurring with $\boldsymbol{G}$, leading to artifacts in the presented image $\boldsymbol{u}$.

3.2.2. With Poisson noise added. We next considered data generated by the system matrix with the addition of Poisson noise. The purpose of this study was to examine the effects of noise on the reconstructions, using data from an otherwise inverse crime case, in which the object model and system model are known exactly. Figure 8 shows images reconstructed from 128 views over the range of r and γ parameters, with profiles plotted in figure 9. Figure 10(a) shows the plot of the CC metric over the range of studied parameters. As in the noiseless case, the CC varied by less than 1.5% across the studied r values for $\gamma > 0.0001$. Unlike the noiseless case, when γ = 0.0001, the CC increased from 0.867 to 0.988 as r increased from 0.0 to 2.0, as the increased blurring provided additional regularity and noise reduction. Noise is also reduced as γ is increased, due to the increased weighting of the TV term. The highest CC value of 0.999 occurred when r = 0.75, the true value of r, and γ = 0.01. As in the noiseless case, contrast and spatial resolution decreased with increasing γ. The FW10M ranged from 12-14 for $\gamma < 1.0$, compared to a true value of 12.

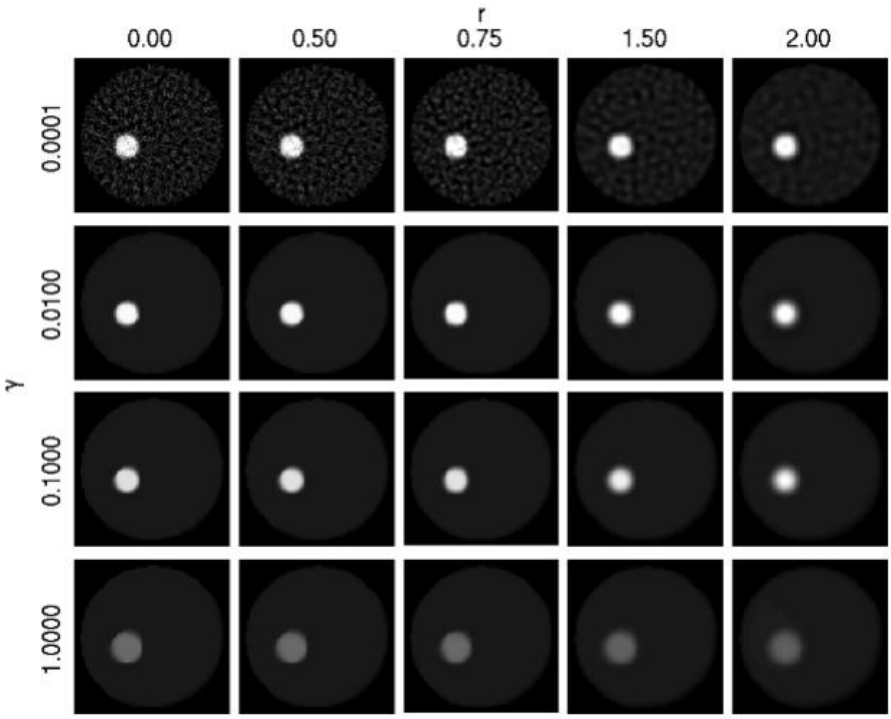


Figure 8. Images reconstructed from 128 views of noisy data using the proposed algorithm with varying values of r and γ. For these images, the projection data were generated by the system matrix.

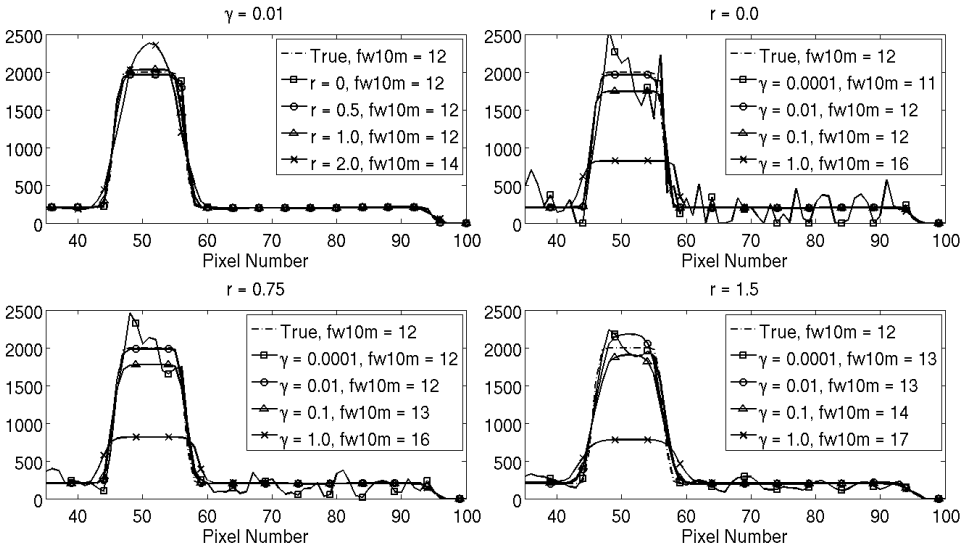


Figure 9. Central diagonal profiles through images reconstructed from noisy inverse crime data from128 views using the proposed algorithm with varying values of r and γ. For these images, the projection data were generated by the system matrix.

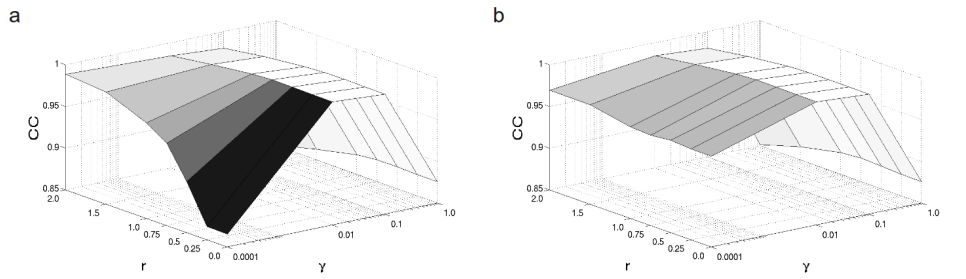


Figure 10. Plots depicting the CC over the range of studied r and γ parameters of images reconstructed from noisy data from 128 views (a) and 9 views (b). For these images, the projection data were generated by the system matrix.

Figures 10(b), 11 and 12 display the images, profiles and plots for noisy images reconstructed from nine views. Similar trends were observed as in the reconstructions from 128 views. Images reconstructed with low γ values ($\gamma = 0.0001$) demonstrated increased noise and streaking artifacts, which were reduced with increasing r.

For $\gamma = 0.01$, the highest CC occurred when the assumed blurring model match the true object ($r = 0.75$), with CC varying by less than 1.4% across the range of r values. The highest CC value of 0.997 was obtained with $\gamma = 0.01$ and $r = 0.75$. Similar FW10M results were obtained using 128 views, with the exception of increased FW10M (14-15) when $\gamma = 0.0001$.

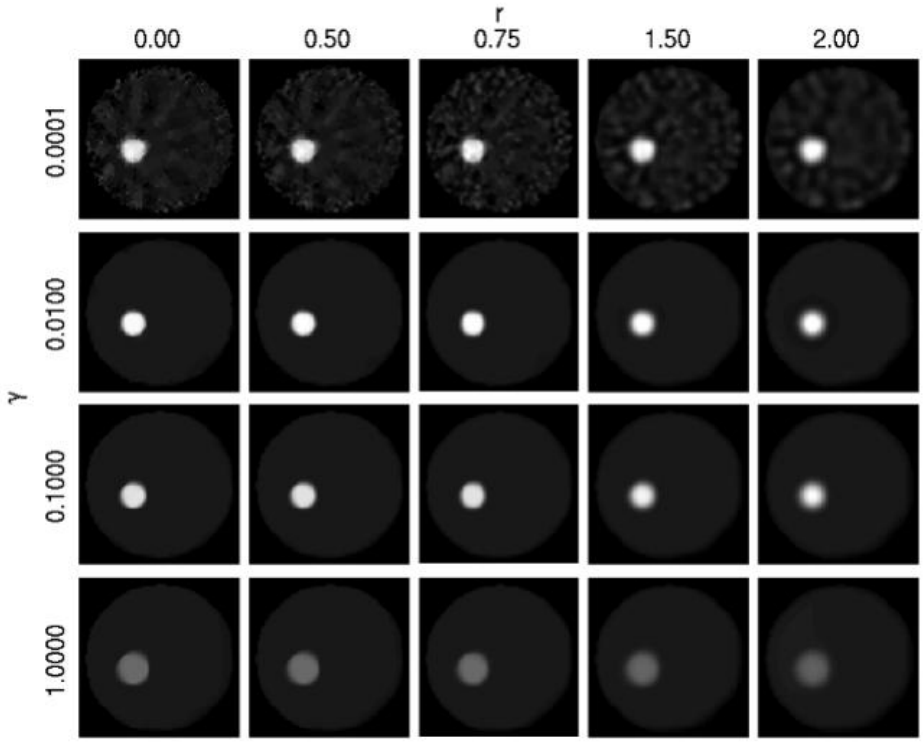


Figure 11. Images reconstructed from 9 views of noisy data using the proposed algorithm with varying values of r and γ. For these images, the projection data were generated by the system matrix.

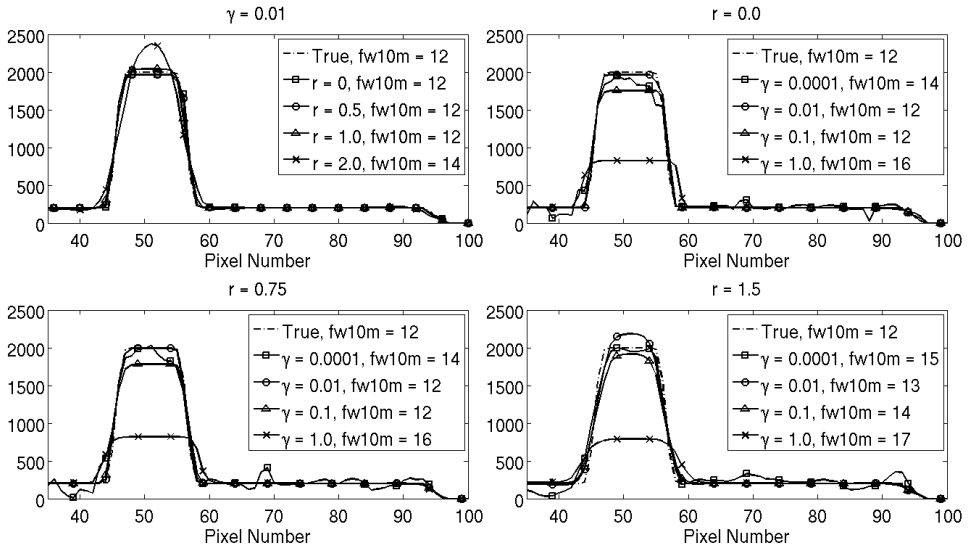


Figure 12. Central diagonal profiles through images reconstructed from noisy inverse crime data from 9 views using the proposed algorithm with varying values of r and γ. For these images, the projection data were generated by the system matrix.

Overall, as in the noiseless case, CC showed little variation with r but greater variation with γ, and both the 9- and 128-view reconstructions suggest that $\gamma = 0.01$ provides the highest CC. The lowest CC value in both the high- and few-view cases occurred with large values of r and γ ($r = 2.0$ and $\gamma = 1.0$).

Figure 13 compares images reconstructed with the proposed reconstruction technique ($\gamma = 0.01$ and $r = 0.75$) and MLEM from data acquired with a varying number of angular views. Table 1 shows CC, SNR and FW10M values for each reconstruction technique and number of views. Images reconstructed using the proposed algorithm had CC values that were 2-4% higher for each case compared to MLEM. The greatest difference is noted for the 9 view case where the image reconstructed using the proposed algorithm yielded a CC of 0.994 while the MLEM image had a CC of 0.954. Streak artifacts were present in the MLEM reconstructions, and were primarily absent in images reconstructed with the proposed algorithm. The noise level in MLEM

reconstructions is higher, leading to lower SNR. Both algorithms provide similar FW10M values compared to the true value of 12.

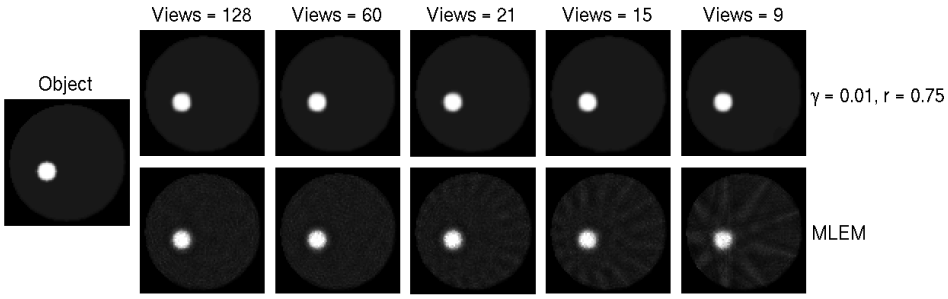


Figure 13. Images reconstructed from noisy projections using the proposed algorithm and MLEM for varying sampling cases. For these images, the projection data were generated by the system matrix.

Table 1. Comparison of image quality metrics from images reconstructed from noisy projections generated by the system matrix.

		128 views	**60 views**	**21 views**	**15 views**	**9 views**
$\gamma = 0.010$ $r = 0.75$	**CC**	0.999	0.998	0.998	0.998	0.999
	SNR	314.7	736.5	154.14	615.16	61.17
	FW10M	12	12	12	12	12
MLEM	**CC**	0.986	0.987	0.984	0.981	0.973
	SNR	6.5	6.13	5.38	6.53	6.31
	FW10M	13	13	14	12	15

4. Monte Carlo simulation study

The purpose of this study was to characterize the performance of the reconstruction technique for the more realistic case where the object does not necessarily match the model assumed in reconstruction, and the modeled system matrix is an approximation to the system that generated the data. In addition, these simulations include realistic effects such as scatter, spatially-varying pinhole sensitivity and blurring from the pinhole aperture.

4.1. Methods

4.1.1. Phantom. The object was defined on a 512 x 512 pixel grid of 0.25 x 0.25 mm pixels. The object consisted of a 28 mm-radius disk of background activity containing five contrast elements of varying size, shape, and intensity, as detailed in Table 2 and displayed in figure 14. Two, two-dimensional Gaussian distributions with peak intensities 638 Bq and 319 Bq and standard deviations 4 mm and 8 mm truncated to have radius 4.4 mm were embedded in the larger disk. Also included in the phantom were a disk representing a cold region with radius 4.4 mm and one disk with radius 2.2 mm having constant intensity, as detailed in Table 2. None of the elements in the phantom were generated by the smoothed piecewise constant model assumed by the reconstruction algorithm, thus representing a challenging reconstruction task.

Table 2. GATE Phantom Specifications.

Element	Radius (mm)	Position (mm)	Intensity (relative)
A	28	(0,0)	Constant; 64 Bq/pixel
B	4.4	(-13,6)	Activity = 6.8 MBq; Peak = 638 Bq/pixel; Std Dev = 4 mm
C	4.4	(6,-13)	Activity = 0.49Mbq; Peak = 319 Bq/pixel; Std Dev = 8 mm
D	4.4	(0,15)	0
E	2.2	(18,4)	Constant; 640 Bq/pixel

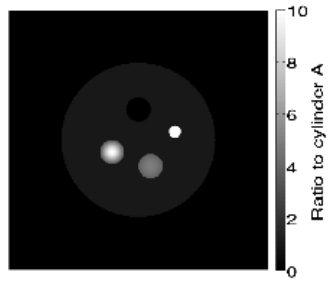


Figure 14. Voxelized phantom used in the GATE studies. The phantom contains contrast elements of varying shape and size as described in Table 2.

4.1.2. Simulations. Projections of the pixelized object were generated using GATE Monte Carlo simulation to model the stochastic emission of photons from a voxelized phantom and their stochastic transmission through the collimator and camera. A three-camera system was simulated. Each collimator was simulated as a 20 mm thick tungsten plate having a 3 mm diameter pinhole with 1.5 mm channel length. A 128 mm x 1 mm NaI crystal was simulated and detected photons binned into 1 mm x 1 mm pixels. Compton scatter, Rayleigh scatter and photoelectric absorption were included as possible interactions for 140 keV photons. Photons detected outside the 129.5 – 150.5 keV range were rejected as scatter. Electronic noise was not modeled. The system is described in figure 15 and table 3.

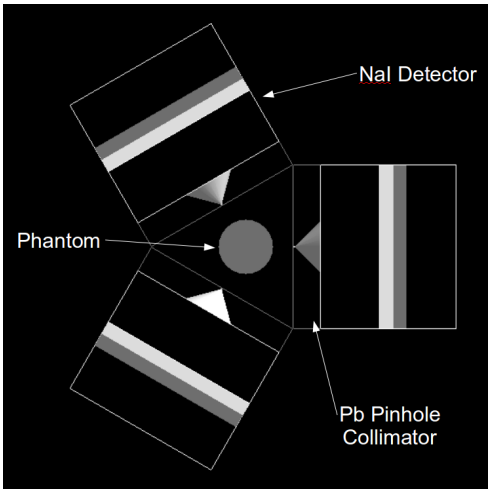


Figure 15. Diagram of Simulated SPECT system.

Table 3. Specifications of the simulated SPECT system.

Camera Size	128 mm x 120 mm x 1 mm
Pinhole diameter	3 mm
Pinhole-to-object distance	35 mm
Pinhole-to-detector distance	63.5 mm

The sensitivity of pinhole collimators depends on the angle of the ray incident on the pinhole. In order to correct for the spatially-varying pinhole sensitivity during reconstruction, a sensitivity map was generated by simulating a flood source on the collimator surface (Vanhove *et al* 2008). The resulting projection represents the spatially-varying sensitivity of the pinhole and was incorporated into the reconstruction algorithm. The sensitivity map was multiplied during each forward projection prior to the summing of data from each ray. Data were multiplied by the sensitivity map prior to backprojection.

Two distinct cases were simulated. In the first case, the total simulated scan time was held constant as the number of views decreased in order to examine the effects of angular undersampling independent of changes in noise. Scans comprising 60, 21, 15, and 9 views distributed over 360 degrees acquired during a 200 second scan were simulated. These data had approximately the same number of total counts in each simulation (~65000 counts). The noise level in SPECT imaging is dependent on the number of detected counts, so the reconstructed images should have similar noise statistics regardless of the number of view angles. The second simulated case held the acquisition time of each view constant across all angular sampling cases. By doing so, the scans that used fewer views had improved temporal sampling, but fewer counts. As the number of views decreased, so did the absolute intensities of the reconstructed images. Images were acquired over 10 seconds for each position of the three-camera gantry, thereby varying the total scan time from 200 seconds for 60 views to 30 seconds for 9 views. In this case, the simulated scan with the fewest views (9) had the fewest counts (~10000 counts) and, consequently, the highest noise level. This represents a more realistic approach for providing dynamic scans with high temporal sampling.

The simulated phantom cannot be described using a constant r across the spatial domain. Each disk has a definite edge and distinct profile. To investigate the effects of varying r in the case where its optimal value is unknown, data were reconstructed using the TV case ($r = 0.0$) and varying r from 0.25 to 2.0 pixels. The TV weighting parameter was varied from 0.0001 to 1.0. The resulting images were evaluated on the basis of reconstruction accuracy with the CC metric as described in section 3.1.3. Spatial resolution was quantified by

considering FW10M of a profile through the center of disk E. A 3 pixel-radius region in the background of disk A was used to calculate SNR. In each case, MLEM reconstructions are also presented as a reference, with the MLEM stopping iteration selected as the iteration with the highest CC value.

4.2. Results

In this section we present the results of the Monte Carlo simulations performed over a range of angular sampling schemes for two different cases: constant total scan time and constant scan time per view.

4.2.1. Constant total scan time. Data reconstructed using the proposed algorithm from 60 views and a variety of r and γ values are presented in figure 15, with profiles presented in figure 17, and plots of CC in figure 18.

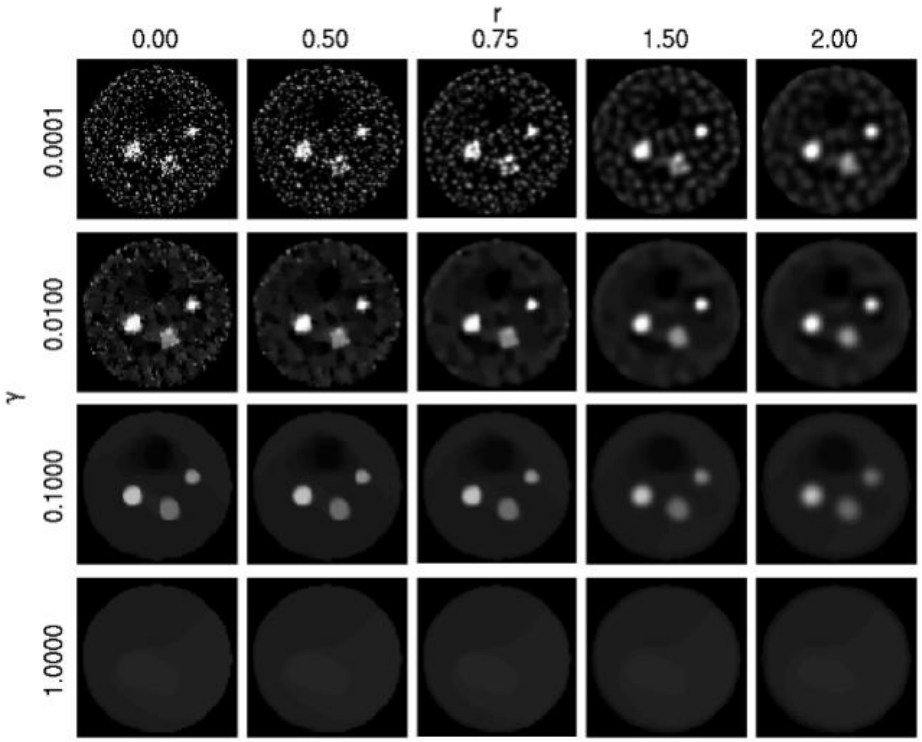


Figure 16. Images reconstructed from 60 views of GATE data simulated for 200 seconds using the proposed algorithm with varying values of r and γ.

When the TV weighting parameter was small ($\gamma = 0.0001$), the resulting image contained high frequency noise; when the TV weighting parameter was large ($\gamma = 1.0$), the object was blurred and contrast reduced. The remainder of the results will focus on $\gamma = 0.1$ and $\gamma = 0.01$. With $\gamma = 0.1$, the CC of the images with the true object varied by 3% across the studied r values, with CC equal to 0.942 at $r = 0.75$ and CC = 0.910 at $r = 2.0$. When $\gamma = 0.1$, the reconstructed profiles do not reach the true peak level for any values of r, as demonstrated in figure 17. Using $\gamma = 0.01$, the profiles reach a higher peak but the CC of these images varies by 11% from 0.946 when $r = 1.5$ to 0.836 at $r = 0.0$. As seen in figure 17, the MLEM reconstructions also did not reach the peak values of the true object profiles, suggesting that this error may be caused by system blurring rather than the

reconstruction algorithm. Using $\gamma = 0.01$, the r value that yielded the optimal image (in terms of CC) was 1.5, compared to an optimal r value of 0.75 when $\gamma = 0.1$. The FW10M for the image reconstructed using $\gamma = 0.01$ and $r = 1.5$ was 7 pixels, compared to a FW10M of 8 pixels resulting from MLEM reconstruction, and a true value of 4. The FW10M for the images reconstructed with $\gamma = 0.1$ and $r = 0.75$ was 8 pixels.

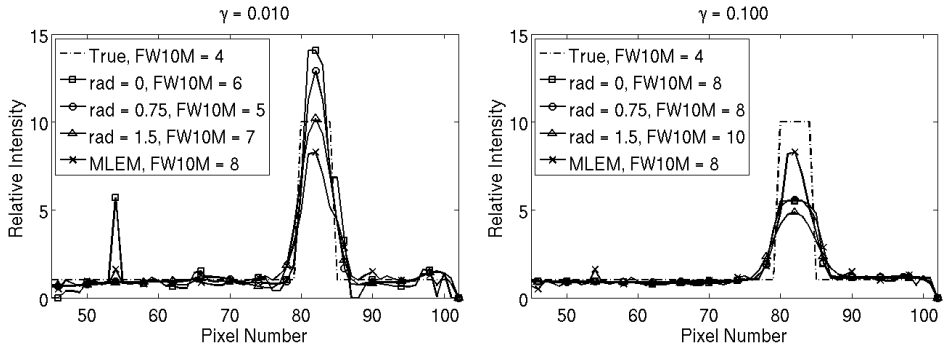


Figure 17. Central vertical profiles through images reconstructed from 60 views of GATE data simulated for 200 seconds using the proposed algorithm with varying values of r and γ.

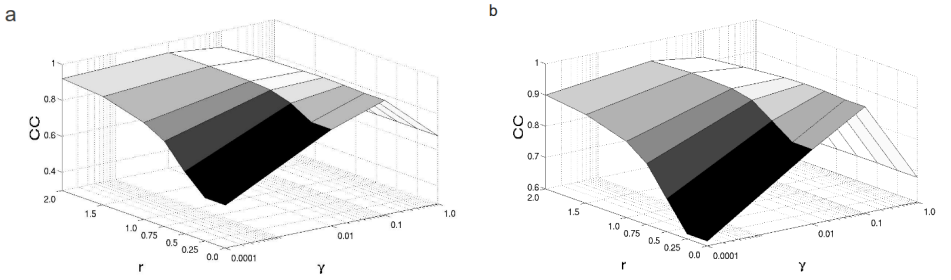


Figure 18. Plots depicting the CC over the range of studied r and γ parameters of images reconstructed from GATE data simulated for 200 seconds, using 60 views (a) and 9 views (b).

The images reconstructed from 9 views demonstrated behavior similar to images reconstructed using 60 views. Images are shown in figure 19. Images reconstructed using $\gamma = 0.01$ contained more noise than images reconstructed using $\gamma = 0.1$. The CC of images reconstructed with $\gamma = 0.01$ vary by 11.3% from 0.946 ($r = 1.5$) to 0.839 ($r = 0.0$), depending on the value of r. Images using $\gamma = 0.1$ had a lesser dependence on r, varying by

3.1% from a 0.945 peak at r = 0.75 to 0.915 at r = 2.0. Images using γ = 0.01 and r = 1.5 have a FW10M value of 8, compared to the true FW10M value of 4 and a FW10M of 8 resulting from MLEM reconstruction. The FW10M of images reconstructed using γ = 0.1 and r = 0.75 was 9 pixels. Profiles are shown in figure 20.

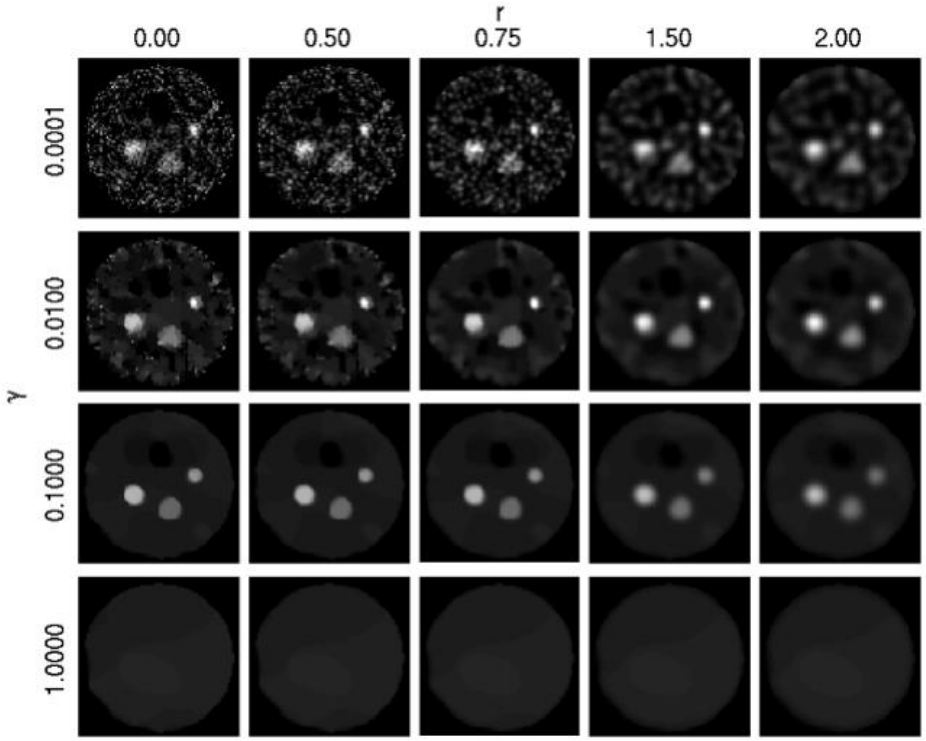


Figure 19. Images reconstructed from 9 views of GATE data simulated for 200 seconds using the proposed algorithm with varying values of r and γ.

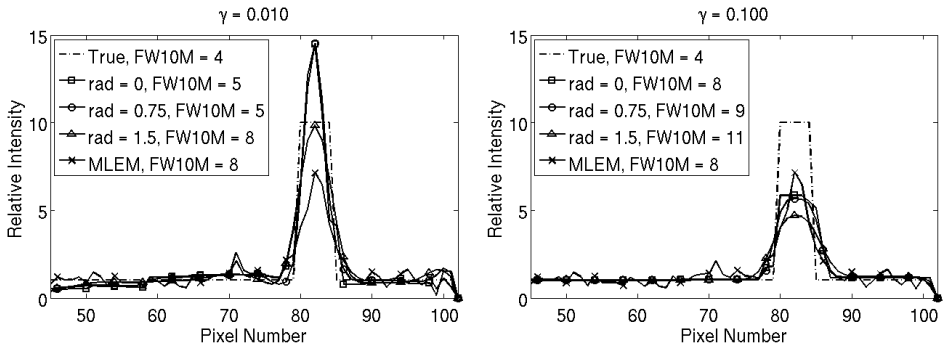


Figure 20. Central vertical profiles through images reconstructed from 9 views GATE data simulated for 200 seconds using the proposed algorithm with varying values of r and γ.

Figure 21 compares images reconstructed with the proposed algorithm ($\gamma = 0.01$, $r = 1.5$ and $\gamma = 0.1$, $r = 0.75$) and MLEM from data acquired with a varying number of angular views. For images reconstructed using the proposed reconstruction technique with $\gamma = 0.01$ and $r = 1.50$, the CC of the images varied by less than 1% from 0.946 to 0.942 as the number of view decreases from 60 to 9. The CC varied similarly for images reconstructed using $\gamma = 0.1$ and $r = 0.75$. For comparison, the CC of images reconstructed by MLEM decreased 6.5% from 0.913 to 0.854 as the number of views degrease from 60 to 9. For this object, the proposed reconstruction algorithm using both $\gamma = 0.01$ and $\gamma = 0.1$ provided higher CC and lower SNR compared to MLEM for all angular cases, while providing similar FW10M values. Images reconstructed using the proposed algorithm contained low-frequency patchy artifacts in the background due to noise, while reducing the streak artifacts present in MLEM reconstructions from few-views.

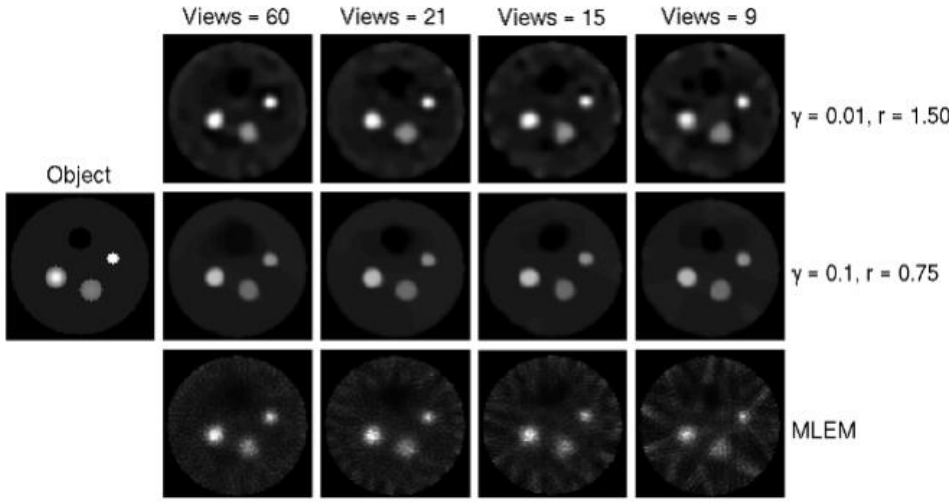


Figure 21. Reconstructions of GATE data simulated for 200s over different numbers of angles using the proposed algorithm and MLEM.

Table 4. Comparison of image quality metrics for images reconstructed from GATE data with the total scan time held constant as the number views decreased.

		60 views	**21 views**	**15 views**	**9 views**
$\gamma = 0.01$, $r = 1.50$	**CC**	0.946	0.945	0.942	0.946
	SNR	17.48	20.57	18.31	6.72
	FW10M	7	8	9	8
$\gamma = 0.10$, $r = 0.75$	**CC**	0.942	0.940	0.941	0.945
	SNR	21.83	24.55	20.95	81.18
	FW10M	8	8	9	9
MLEM	**CC**	0.913	0.901	0.889	0.854
	SNR	4.94	4.92	3.65	4.15
	FW10M	8	11	9	8

4.2.2. Constant scan time per view. This set of simulations modeled a constant scan time per view (i.e., decreasing total scan time with decreasing number of views), representing the case where temporal sampling improves as the number of views decreases. Figures 22-24 present images reconstructed from 9 views with 10 seconds per view (compared to 66.67 seconds per view in figures 17b, 19 and 20).

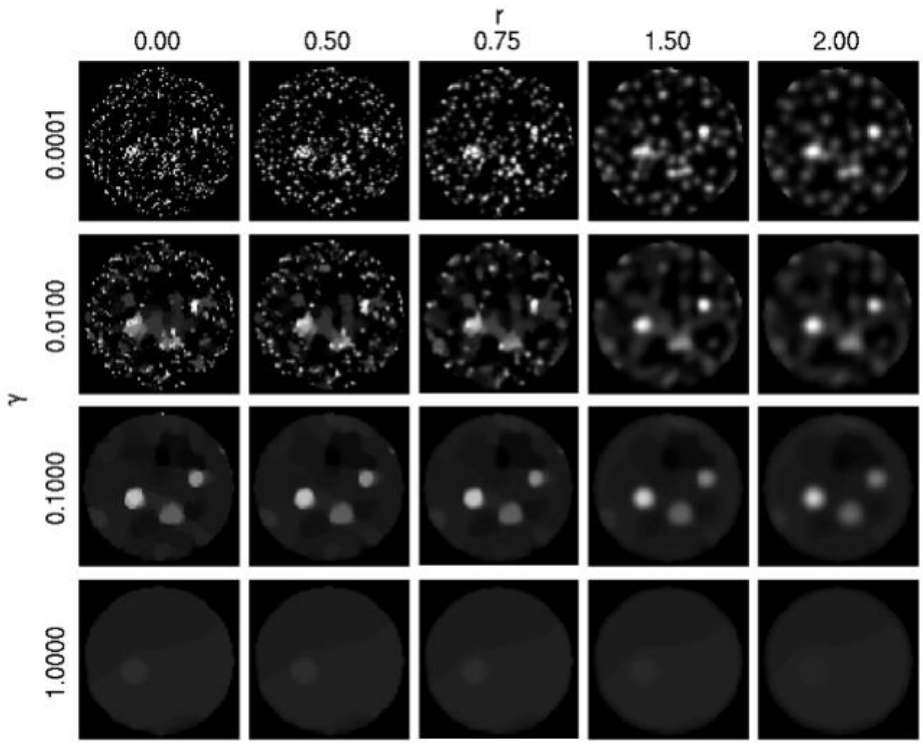


Figure 22. Images reconstructed from 9 views of GATE data simulated for 30 seconds using the proposed algorithm with varying values of r and γ.

Images reconstructed from nine views using $\gamma = 0.01$ had lower reconstruction accuracy (CC < 0.9) compared to the images reconstructed from a 200 second scan time presented in the previous section. Using $\gamma = 0.1$, a maximum CC value of 0.921 occurred when $r = 0.75$. Similar to the 200 second scans, the CC varied by less than 2% across the range of r values for $\gamma = 0.1$. However, unlike the 200 second scans, the 30 second scans showed a larger variation in CC (~40%) across the range of r values for γ not equal to 0.1. In addition to increasing CC, using $\gamma = 0.1$ resulted in reduced noise but increased blurring (higher FW10M) compared to $\gamma = 0.01$.

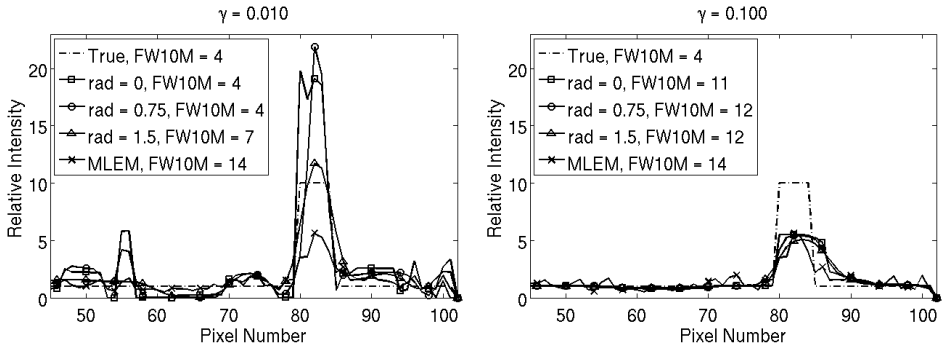


Figure 23. Central vertical profiles through images reconstructed from 9 views GATE data simulated for 30 seconds using the proposed algorithm with varying values of r and γ.

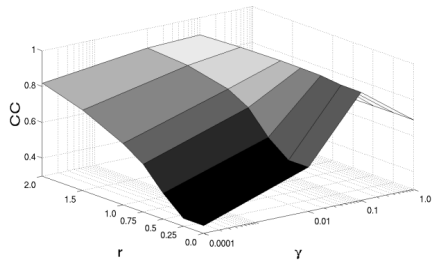


Figure 24. Plots depicting the CC over the range of studied r and γ parameters of images reconstructed from GATE data simulated for 9 views over 30 seconds.

Figure 25 compares images reconstructed with the proposed algorithm (γ = 0.01, r = 1.5 and γ = 0.01, r = 0.75) and MLEM from data acquired with a varying number of angular views (9 to 60) and a constant 10 second acquisition time for each angular position of the three-camera system. Thus the total scan time was 200, 70, 50, and 30 seconds for 60, 21, 15, and 9 views, respectively. Associated image quality metrics are presented in Table 5. As scan time and angular sampling decreased, images reconstructed using the proposed algorithm with γ = 0.01 and r = 1.50 show decreased accuracy compared to scans with less noise and the same angular sampling presented in the previous section. When reconstructing from 21 views, 15 views and 9 views, higher CC is achieved using γ = 0.1 and r = 0.75, compared to using γ = 0.01 and r = 1.50. In both cases, the proposed reconstruction algorithm provides higher CC and SNR than MLEM. For reconstructions using γ = 0.01 and r =

1.5, CC varied by 7.5% from 0.946 to 0.875 as the number of views was reduced from 60 to 9. A variation of only 2.6% from 0.942 to 0.921 was measured for images reconstructed using $\gamma = 0.1$ and $r = 0.75$. These reductions were both less than the decrease measured for MLEM, which saw a decrease of 12.8%, from 0.915 to 0.798 as the number of views was reduced from 60 to 9.

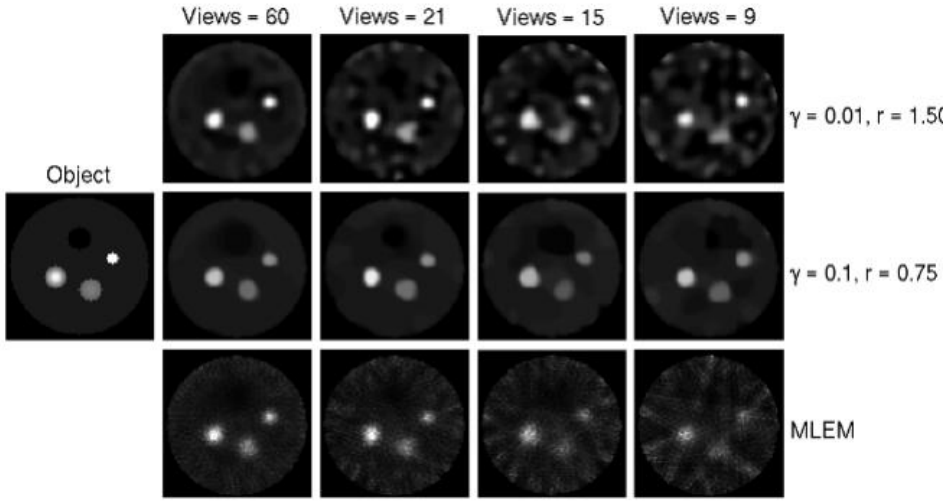


Figure 25. Images reconstructed using the proposed algorithm and MLEM from GATE data simulated with the same time per view for different numbers of views.

Table 5. Comparison of image quality metrics for images reconstructed from GATE data with varying number of views and constant scan time per view.

		60 views	**21 views**	**15 views**	**9 views**
$\gamma = 0.01$, $r = 1.50$	**CC**	0.946	0.915	0.889	0.875
	SNR	17.48	9.96	1.29	10.82
	FW10M	7	6	9	7
$\gamma = 0.10$, $r = 0.75$	**CC**	0.942	0.937	0.908	0.921
	SNR	21.83	5149.86	24.76	52.09
	FW10M	8	8	10	12
MLEM	**CC**	0.915	0.885	0.829	0.798
	SNR	4.94	2.98	3.49	4.06
	FW10M	8	9	8	14

5. Discussion

The presented simulations investigated the proposed reconstruction technique over a range of objects, noise conditions, and angular sampling schemes. Overall, the results demonstrate that blurring and noise regularization increased with increasing values of r, the standard deviation of the Gaussian blurring kernel, and γ, the TV weighting parameter. For example, in the high-view case with noiseless data generated from the system model, reconstructions using the lowest γ value studied ($\gamma = 0.0001$) yielded the most accurate images for a given value of r. When the data were made inconsistent by the addition of Poisson noise, the optimal studied γ value increased to $\gamma = 0.01$. These cases, in which data generated using the system matrix and the blurring model were used in reconstruction, indicate that accurate reconstruction is possible when the incorrect blurring model is used, as there was only a 2.5% decrease in the CC metric over all r studied when $\gamma = 0.1$ and $\gamma = 0.01$. However, in the few-view case, using an r larger than r_{true} caused the number of meaningful coefficients in the intermediate image f to increase rapidly compared to using lower values of r. This indicates a less sparse image, limiting the effectiveness of exploiting gradient-magnitude sparsity to reduce the number of views needed for reconstruction. When an approximately accurate blurring model is used, the intermediate image f is the most sparse in the gradient-magnitude sense. This may allow a greater reduction in the sampling necessary for reconstruction.

When noisy data were generated using GATE Monte Carlo simulations, larger r had a benefit when lower γ were used. For instance, in the 9 view case when data were simulated for 200 seconds and images were reconstructed with $\gamma = 0.01$, the CC varied by 11% over the range of studied r values, with a high r value ($r = 1.5$) yielding the most accurate reconstructions. When $\gamma = 0.1$ was used, a lower r ($r = 0.75$) yielded the most accurate reconstructions. Similarly, when the scan time was decreased in the few-view case, $r = 2.0$ yielded the most accurate reconstructions when $\gamma = 0.01$ was used; however, the highest overall CC in the few-view, decreasing scan time case was obtained with $\gamma = 0.1$ and $r = 0.75$. Reconstructions using both $\gamma = 0.01$, $r = 2.0$

and $\gamma = 0.1$, $r = 0.75$ have similar CC but different qualitative attributes (figures 21 and 25). The preferred parameter combination requires further study with observers. Overall, reconstructions from data generated using GATE simulations suggest that when the true blurring model is unknown and noise is present, lower values of γ ($\gamma = 0.01$ in this particular study) benefit from larger r values, while $\gamma = 0.1$ benefits from lower r values, with a smaller dependence on r. Since the inverse crime study demonstrated that smaller r values result in a more sparse intermediate image, the combination of $\gamma = 0.1$ and $r = 0.75$ may be advantageous for reconstruction from few-views.

The results also suggest that, when an appropriate value of the TV penalty term is included in the proposed reconstruction algorithm ($\gamma = 0.01$ or 0.1 for the cases studied), streaking artifacts are reduced compared to MLEM reconstructions. While images reconstructed with the proposed algorithm contain higher SNR, low frequency variations (patchy artifacts) were seen in high-noise simulation cases (figures 22 and 25). Low frequency, patchy artifacts have been noted in CT TV reconstructions from noisy data, and future work is required to quantify the impact of these artifacts on the ability of observers to identify objects of diagnostic interest. (Tang *et al* 2009).

The presented work suggests potential benefits of the proposed reconstruction algorithm compared to MLEM, however, additional work is required for a systematic comparison, including experimental investigation. One limitation of the presented work is that the simulations modeled 2D objects and acquisition, whereas SPECT data are acquired in three dimensions. We hypothesize that the principles and model presented in this work can be generalized to a 3D case with the expansion of the system matrix and applying the blurring-masking function in three dimensions. Additional studies are necessary to investigate this hypothesis. Reconstruction from multi-pinhole systems could be accomplished by modifying the system matrix to include contributions from all pinholes. Future work is also planned to apply the reconstruction technique to *in vivo* data to investigate the assumption that SPECT objects may be modeled as blurred piecewise constant objects. Future work will also investigate the performance of this algorithm for dynamic imaging from few-views.

6. Conclusions

This study proposed and characterized a sparsity-exploiting reconstruction algorithm for SPECT that is intended for few-view imaging and that phenomenologically models the object as piecewise constant subject to a blurring operation. While the reconstruction technique assumes a specific blurring model, the results demonstrate that the knowledge of the true blurring parameter is not required for accurate reconstruction, as the reconstruction algorithm has limited sensitivity to r in the low noise cases and benefits from increasing r in the high noise case. However, the results suggest that accurately modeling the blurring parameter provides increased gradient-magnitude sparsity, which may enable further reductions in sampling. The reconstructed images demonstrate that the reconstruction algorithm introduces low-frequency artifacts in the presence of noise, but eliminates streak artifacts due to angular undersampling. The effects of these artifacts on observers will be studied in future work. Overall, the results demonstrate preliminary feasibility of a sparsity-exploiting reconstruction algorithm which may be beneficial for few-view SPECT.

Acknowledgments

This work was supported in part by NIH R15 grant CA143713 and R01 grants CA158446, CA120540 and EB000225. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health. This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. The high performance computing resources used in this paper were funded by NSF grant OCI-0923037.

References

Barrett H H and Myers K J 2004 *Foundations of Image Science* ed E A Saleh, Bahaa (Hoboken, NJ: John Wiley & Sons, Inc.)

Beekman F J, van der Have F, Vastenhouw B, van der Linden A J a, van Rijk P P, Burbach J P H and Smidt M P 2005 U-SPECT-I: a novel system for submillimeter-resolution tomography with radiolabeled molecules in mice. *J. Nucl. Med.* **46** 1194–200

Beekman F J and Vastenhouw B 2004 Design and simulation of a high-resolution stationary SPECT system for small animals *Phys. Med. Biol.* **49** 4579–92

Chambolle A and Pock T 2011 A first-order primal-dual algorithm for convex problems with applications to imaging *J. Math. Imag. Vis.* **40** 1–26

Chen G-H, Tang J and Leng S 2008 Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets *Med. Phys.* **35** 660–3

Donato P, Coelho P, Santos C, Bernardes A and Caseiro-Alves F 2012 Correspondence between left ventricular 17 myocardial segments and coronary anatomy obtained by multi-detector computed tomography: an ex vivo contribution. *Surgical and Radiologic Anatomy* **34** 805–10

Duan X, Zhang L, Xing Y, Chen Z and Cheng J 2009 Few-View Projection Reconstruction With an Iterative Reconstruction-Reprojection Algorithm and TV Constraint *IEEE Trans. Nucl. Sci.* **56** 1377–82

Furenlid L R, Wilson D W, Chen Y-C, Kim H, Pietraski P J, Crawford M J and Barrett H H 2004 FastSPECT II: A Second-Generation High-Resolution Dynamic SPECT Imager. *IEEE Trans. Nucl. Sci.* **51** 631–5

Gullberg G T 2004 Dynamic SPECT imaging: exploring a new frontier in medical imaging *IEEE International Symposium on Biomedical Imaging: Nano to Macro* 607–10

Gullberg G T, Reutter B W, Sitek A, Maltz J S and Budinger T F 2010 Dynamic single photon emission computed tomography--basic principles and cardiac applications. *Phys. Med. Biol.* **55** R111–91

Hill D L, Batchelor P G, Holden M and Hawkes D J 2001 Medical image registration *Phys. Med. Biol.* **46** R1–45

Hudson H and Larkin R 1994 Accelerated image reconstruction using ordered subsets of projection data *IEEE Trans. Med. Img.* **13** 601–9

Jain R 1988 Determinants of tumor blood flow: a review *Cancer Research* **48** 2641–58

Jan S, Santin G, Strul D and Staelens S 2004 GATE: a simulation toolkit for PET and SPECT *Phys. Med. Biol.* **49** 4543

Kaipio J and Somersalo E 2005 *Statistical and Computational Inverse Problems* ed S S Antman, J E Marsden and L Sirovich (New York: Springer Science+Business Media, LLC)

Ma D, Wolf P, Clough A and Schmidt T 2012 The Performance of MLEM for Dynamic Imaging From Simulated Few-View, Multi-Pinhole SPECT *IEEE Trans. Nucl. Sci*

Pereztol-Valdés O, Candell-Riera J, Santana-Boado C, Angel J, Aguadé-Bruix S, Castell-Conesa J, Garcia E V and Soler-Soler J 2005 Correspondence between left ventricular 17 myocardial segments and coronary arteries. *European Heart Journal* **26** 2637–43

Ritschl L, Bergner F, Fleischmann C and Kachelriess M 2011 Improved total variation-based CT image reconstruction applied to clinical data. *Phys. Med. Biol.* **56** 1545–61

Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Img* **1** 113–22Siddon R L 1985 Fast calculation of the exact radiological path for a three-dimensional CT array *Med. Phys.* **12** 252–5

Sidky E, Kao C and Pan X 2006 Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT *J. X-ray Sci. Tech.* **14** 119–39

Sidky E Y, Jørgensen J H and Pan X 2012 Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm *Phys. Med. Biol.* **57** 3065–91

Sidky E Y and Pan X 2008 Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys. Med. Biol.* **53** 4777–807

Sidky E Y, Pan X, Reiser I S, Nishikawa R M, Moore R H and Kopans D B 2009 Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms *Med. Phys.* **36** 4920–32

Tang J, Nett B E and Chen G-H 2009 Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms. *Phys. Med. Biol.* **54** 5781–804

Vandenberghe S, D'Asseler Y, Van de Walle R, Kauppinen T, Koole M, Bouwens L, Van Laere K, Lemahieu I and Dierckx R a 2001 Iterative reconstruction algorithms in nuclear medicine. *Comput. Med. Imaging Graphics* **25** 105–11

Vanhove C, Defrise M, Lahoutte T and Bossuyt A 2008 Three-pinhole collimator to improve axial spatial resolution and sensitivity in pinhole SPECT. *Eur. J. Nucl. Med. Mol. Imaging* **35** 407–15

Appendix E

Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm

Phys. Med. Biol., vol. 57, issue 10, pp. 3065–3091, 2012.
doi:10.1088/0031-9155/57/10/3065.
Published 27 April 2012.

E. Y. Sidky, **J. H. Jørgensen** and X. Pan

IOP Publishing Physics in Medicine and Biology
Phys. Med. Biol. **57** (2012) 3065–3091 doi:10.1088/0031-9155/57/10/3065

Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm

Emil Y Sidky[1], **Jakob H Jørgensen**[2] **and Xiaochuan Pan**[1]

[1] Department of Radiology, University of Chicago, 5841 S. Maryland Ave., Chicago, IL 60637, USA
[2] Department of Informatics and Mathematical Modeling, Technical University of Denmark, Richard Petersens Plads, Building 321, 2800 Kgs. Lyngby, Denmark

E-mail: sidky@uchicago.edu, jakj@imm.dtu.dk and xpan@uchicago.edu

Received 23 November 2011, in final form 6 March 2012
Published 27 April 2012
Online at stacks.iop.org/PMB/57/3065

Abstract
The primal–dual optimization algorithm developed in Chambolle and Pock (CP) (2011 *J. Math. Imag. Vis.* **40** 1–26) is applied to various convex optimization problems of interest in computed tomography (CT) image reconstruction. This algorithm allows for rapid prototyping of optimization problems for the purpose of designing iterative image reconstruction algorithms for CT. The primal–dual algorithm is briefly summarized in this paper, and its potential for prototyping is demonstrated by explicitly deriving CP algorithm instances for many optimization problems relevant to CT. An example application modeling breast CT with low-intensity x-ray illumination is presented.

(Some figures may appear in colour only in the online journal)

1. Introduction

Optimization-based image reconstruction algorithms for CT have been investigated heavily recently due to their potential to allow for reduced scanning effort while maintaining or improving image quality (McCollough *et al* 2009, Pan *et al* 2009). Such methods have been considered for many years, but during the past five years computational barriers have been lowered enough such that iterative image reconstruction can be considered for practical application in CT (Ziegler *et al* 2008). The transition to practice has been taking place alongside further theoretical development particularly with algorithms based on the sparsity-motivated ℓ_1-norm (Li *et al* 2002, Sidky *et al* 2006, 2010, Sidky and Pan 2008, Chen *et al* 2008, Ritschl *et al* 2011, Defrise *et al* 2011, Ramani and Fessler 2011, Jørgensen *et al* 2011a). Despite the recent interest in sparsity, optimization-based image reconstruction algorithm development continues to proceed along many fronts and there is as of yet no consensus on a particular

optimization problem for the CT system. In fact, it is beginning to look like the optimization problem, upon which the iterative image reconstruction algorithms are based, will themselves be subject to design depending on the particular properties of each scanner type and imaging task.

Considering the possibility of tailoring optimization problems to a class of CT scanners makes the design of iterative image reconstruction algorithms a daunting task. Optimization formulations generally construct an objective function comprised of a data fidelity term and possible penalty terms discouraging unphysical behavior in the reconstructed image, and they possibly include hard constraints on the image. The image estimate is arrived at by extremizing the objective subject to any constraints placed on the estimate. The optimization problems for image reconstruction can take many forms depending on image representation, projection model and objective and constraint design. On top of this, it is difficult to solve many of the optimization problems of interest. A change in optimization problem formulation can mean many weeks or months of algorithm development to account for the modification.

Due to this complexity, it would be quite desirable to have an algorithmic tool to facilitate design of optimization problems for CT image reconstruction. This tool would consist of a well-defined set of mechanical steps that generate a convergent algorithm from a specific optimization problem for CT image reconstruction. The goal of this tool would be to allow for rapid prototyping of various optimization formulations; one could design the optimization problem free of any restrictions imposed by a lack of an algorithm to solve it. The resulting algorithm might not be the most efficient solver for the particular optimization problem, but it would be guaranteed to give the answer.

In this paper we consider convex optimization problems for CT image reconstruction, including non-smooth objectives, unconstrained and constrained formulations. One general algorithmic tool is to use steepest descent or projected steepest descent (Nocedal and Wright 2006). Such algorithms, however, do not address non-smooth objective functions and they have difficulty with constrained optimization, being applicable for only simple constraints such as non-negativity. Another general strategy involves some form of evolving quadratic approximation to the objective. The literature on this flavor of algorithm design is enormous, including nonlinear conjugate gradient (CG) methods (Nocedal and Wright 2006), parabolic surrogates (Erdogan and Fessler 1999, Defrise *et al* 2011) and iteratively reweighted least-squares (Green 1984). For the CT system, these strategies often require quite a bit of know-how due to the very large scale and ill-posedness of the imaging model. Once the optimization formulation is established, however, these quadratic methods provide a good option to gain in efficiency.

One of the main barriers to prototyping alternative optimization problems for CT image reconstruction is the size of the imaging model; volumes can contain millions of voxels and the sinogram data can correspondingly consist of millions of x-ray transmission measurements. For large-scale systems, there has been some resurgence of first-order methods (Yin *et al* 2008, Combettes and Pesquet 2008, Beck and Teboulle 2009, Becker *et al* 2010, Chambolle and Pock 2011, Jensen *et al* 2011) and recently there has been applications of first-order methods specifically for optimization-based image reconstruction in CT (Jensen *et al* 2011, Choi *et al* 2010, Jørgensen *et al* 2011b). These methods are interesting because they can be adapted to a wide range of optimization problems involving non-smooth functions such as those involving ℓ_1-based norms. In particular, the algorithm that we pursue further in this paper is a first-order primal–dual algorithm for convex problems by Chambolle and Pock (2011). This algorithm goes a long way toward the goal of optimization problem prototyping because it covers a very general class of optimization problems that contain many optimization formulations of interest to the CT community.

For a selection of optimization problems of relevance to CT image reconstruction, we work through the details of setting up the Chambolle–Pock (CP) algorithm. We refer to these dedicated algorithms as *algorithm instances*. Our numerical results demonstrate that the algorithm instances achieve the solution of difficult convex optimization problems under challenging conditions in reasonable time and without parameter tuning. In section 2, the CP methodology and algorithm is summarized; in section 3, various optimization problems for CT image reconstruction are presented along with their corresponding CP algorithm instance and section 4 shows a limited study on a breast CT simulation that demonstrates the application of the derived CP algorithm instances.

2. Summary of the generic CP algorithm

The Chambolle and Pock (2011) (CP) algorithm is primal–dual meaning that it solves an optimization problem simultaneously with its dual. On its face, it would seem to involve extra work by solving two problems instead of one, but the algorithm comes with convergence guarantee and solving both problems provides a robust, non-heuristic convergence check—the duality gap.

The CP algorithm applies to a general form of the primal minimization:

$$\min_x \{F(Kx) + G(x)\}, \tag{1}$$

and a dual maximization:

$$\max_y \{-F^*(y) - G^*(-K^T y)\}, \tag{2}$$

where x and y are finite-dimensional vectors in the spaces X and Y, respectively; K is a linear transform from X to Y; G and F are convex, possibly non-smooth, functions mapping the respective X and Y spaces to non-negative real numbers and the superscript '*' in the dual maximization problem refers to convex conjugation, defined in equations (3) and (4). We note that the matrix K need not be a square; X and Y will in general have different dimensions. Given a convex function H of a vector $z \in Z$, its conjugate can be computed by the Legendre transform (Rockafellar *et al* 1970), and the original function can be recovered by applying conjugation again:

$$H^*(z) = \max_{z'} \{\langle z, z' \rangle_Z - H(z')\}, \tag{3}$$

$$H(z') = \max_{z} \{\langle z', z \rangle_Z - H^*(z)\}. \tag{4}$$

The notation $\langle \cdot, \cdot \rangle_Z$ refers to the inner product in the vector space Z.

Formally, the primal and dual problems are connected in a generic saddle point optimization problem:

$$\min_x \max_y \{\langle Kx, y \rangle_Y + G(x) - F^*(y)\}. \tag{5}$$

By performing the maximization over y in equation (5), using equation (4) with Kx associated with y', the primal minimization (1) is derived. Similarly, performing the minimization over x in equation (5), using equation (3) and the identity $\langle Kx, y \rangle = \langle x, K^T y \rangle$, yields the dual maximization (2), where the T superscript denotes matrix transposition.

The minimization problem in equation (1), though compact, covers many minimization problems of interest to tomographic image reconstruction. Solving the dual problem, equation (2), simultaneously allows for the assessment of algorithm convergence. For intermediate estimates x and y of the primal minimization and the dual maximization, respectively, the primal objective will be greater than or equal to the dual objective. The

difference between these objectives is referred to as the duality gap, and convergence is achieved when this gap is zero. Plenty of examples of useful optimization problems for tomographic image reconstruction will be described in detail in section 3, but first we summarize algorithm 1 from Chambolle and Pock (2011).

Algorithm 1. Pseudocode for N steps of the basic Chambolle–Pock algorithm. The constant L is the ℓ_2-norm of the matrix K; τ and σ are non-negative CP algorithm parameters, which are both set to $1/L$ in the present application; $\theta \in [0, 1]$ is another CP algorithm parameter, which is set to 1; and n is the iteration index. The proximal operators $prox_\sigma$ and $prox_\tau$ are defined in equation (6).

$L \leftarrow \|K\|_2;\ \tau \leftarrow 1/L;\ \sigma \leftarrow 1/L;\ \theta \leftarrow 1;\ n \leftarrow 0$
initialize x_0 and y_0 to zero values
$\bar{x}_0 \leftarrow x_0$
repeat
  $y_{n+1} \leftarrow prox_\sigma[F^*](y_n + \sigma K \bar{x}_n)$
  $x_{n+1} \leftarrow prox_\tau[G](x_n - \tau K^T y_{n+1})$
  $\bar{x}_{n+1} \leftarrow x_{n+1} + \theta(x_{n+1} - x_n)$
  $n \leftarrow n + 1$
until $n \geqslant N$

2.1. Chambolle–Pock: algorithm 1

The CP algorithm simultaneously solves equations (1) and (2). As presented in Chambolle and Pock (2011), the algorithm is simple, yet extremely effective. We repeat the steps here in listing 1 for completeness, providing the parameters that we use for all results shown below. The parameter descriptions are provided in Chambolle and Pock (2011), but note that in our usage specified above there are *no* free parameters. This is an extremely important feature for our purpose of optimization prototyping. One caveat is that technically the proof of convergence for the CP algorithm assumes $L^2\sigma\tau < 1$, but in practice we have never encountered a case where the choice $\sigma = \tau = 1/L$ failed to tend to convergence. We stress that in equation (2) the matrix K^T needs to be the transpose of the matrix K; this point can sometimes be confusing because K for imaging applications is often intended to be an approximation to some continuous operator such as projection or differentiation and often K^T is taken to mean the approximation to the continuous operator's adjoint, which may or may not be the matrix transpose of K. The constant L is the magnitude of the matrix K, its largest singular value. Appendix A gives the details on computing L via the power method. The key to deriving the particular algorithm instances are the proximal mappings $prox_\sigma[F^*]$ and $prox_\tau[G]$ (called resolvent operators in Chambolle and Pock (2011)).

The proximal mapping is used to generate a descent direction for the convex function H and it is obtained by the following minimization:

$$prox_\sigma[H](z) = \arg\min_{z'} \left\{ H(z') + \frac{\|z - z'\|_2^2}{2\sigma} \right\}. \tag{6}$$

This operation does admit non-smooth convex functions, but H does need to be simple enough that the above minimization can be solved in a closed form. For CT applications, the ability to handle non-smooth F and G allows the study of many optimization problems of recent interest, and the simplicity limitation is not that restrictive as will be seen.

2.2. The CP algorithm for prototyping of convex optimization problems

To prototype a particular convex optimization problem for CT image reconstruction with the CP algorithm, there are five basic steps.

(1) Map the optimization problem to the generic minimization problem in equation (1).

(2) Derive the dual maximization problem, equation (2), by computing the convex conjugates of F and G using the Legendre transform (3).

(3) Derive the proximal mappings of F^* and G using equation (6).

(4) Substitute the results of (3) into the generic CP algorithm in listing 1 to obtain a CP algorithm instance.

(5) Run the algorithm, monitoring the primal–dual gap for convergence.

As will be seen below, a great variety of constrained and unconstrained optimization problems can be written in the form of equation (1). Specifically, using the algebra of convex functions (Rockafellar *et al* 1970), that the sum of two convex functions is convex and that the composition of a convex function with a linear transform is a convex function, many interesting optimization formulations can be put in the form of equation (1). We will also make use of convex functions which are not smooth—notably ℓ_1 based norms and indicator functions $\delta_S(x)$:

$$\delta_S(x) \equiv \begin{cases} 0 & x \in S \\ \infty & x \notin S \end{cases}, \tag{7}$$

where S is a convex set. The indicator function is particularly handy for imposing constraints. In computing the convex conjugate and proximal mapping of convex functions, we make much use of the standard calculus rule for extremization, $\nabla f = 0$, but such computations are also augmented with geometric reasoning, which may be unfamiliar. Accordingly, we have included appendices to show some of these computation steps. With this quick introduction, we are now in a position to derive various algorithm instances for CT image reconstruction from different convex optimization problems.

3. CP algorithm instances for CT

For this paper, we only consider optimization problems involving the linear imaging model for x-ray projection, where the data are considered as line integrals over the object's x-ray attenuation coefficient. Generically, maintaining consistent notation with Chambolle and Pock (2011), the discrete-to-discrete CT system model (Barrett and Myers 2004) can be written as

$$Au = g, \tag{8}$$

where A is the projection matrix taking an object represented by expansion coefficients u and generating a set of line-integration values g. This model covers a multitude of expansion functions and CT configurations, including both 2D fan-beam and 3D cone-beam projection data models.

A few notes on notation are in order. In the following, we largely avoid indexing of the various vector spaces in order that the equations and pseudocode listings are brief and clear. Any of the standard algebraic operations between vectors is to be interpreted in a componentwise manner unless explicitly stated. Also an algebraic operation between a scalar and a vector is to be distributed among all components of the vector, e.g. $1 + v$ adds one to all components of v. For the optimization problems below, we employ three vector spaces: I the space of discrete images in either 2 or 3 dimensions; D the space of the CT sinograms (or projection data); and V the space of spatial-vector-valued image arrays, $V = I^d$, where $d = 2$ or 3 for 2D and 3D-space, respectively. For the CT system model (equation (8)), $u \in I$

and $g \in D$, but we note that the space D can also include sinograms which are not consistent with the linear system matrix A. The vector space V will be used below for forming the total variation (TV) semi-norm; an example of such a vector $v \in V$ is the spatial gradient of an image u. Although the pixel representation is used, much of the following can be applied to other image expansion functions. As we will be making much use of certain indicator functions, we define two important sets, $Box(a)$ and $Ball(a)$, through their indicator function:

$$\delta_{\mathrm{Box}(a)}(x) \equiv \begin{cases} 0 & \|x\|_\infty \leqslant a \\ \infty & \|x\|_\infty > a \end{cases} \tag{9}$$

and

$$\delta_{\mathrm{Ball}(a)}(x) \equiv \begin{cases} 0 & \|x\|_2 \leqslant a \\ \infty & \|x\|_2 > a \end{cases}. \tag{10}$$

Recall that the $\|\cdot\|_\infty$ norm selects the largest component of the argument; thus, $Box(a)$ comprises vectors with no component larger than a (in 2D $Box(a)$ is a square centered on the origin with width $2a$). We also employ $\mathbf{0}_X$ and $\mathbf{1}_X$ to mean a vector from the space X with all components set to 0 and 1, respectively.

3.1. Image reconstruction by least-squares

Perhaps the simplest optimization method for performing image reconstruction is to minimize the quadratic data error function. We present this familiar case in order to gain some experience with the mechanics of deriving CP algorithm instances, and because the quadratic data error term will play a role in other optimization problems below. The primal problem of interest is

$$\min_u \tfrac{1}{2}\|Au - g\|_2^2. \tag{11}$$

To derive the CP algorithm instance, we make the following mechanical associations with the primal problem (1):

$$F(y) = \tfrac{1}{2}\|y - g\|_2^2, \tag{12}$$

$$G(x) = 0, \tag{13}$$

$$x = u, \quad y = Au \tag{14}$$

$$K = A. \tag{15}$$

Applying equation (3), we obtain the convex conjugates of F and G:

$$F^*(p) = \tfrac{1}{2}\|p\|_2^2 + \langle p, g\rangle_D, \tag{16}$$

$$G^*(q) = \delta_{\mathbf{0}_I}(q), \tag{17}$$

where $p \in D$ and $q \in I$. While obtaining F^* in this case involves elementary calculus for extremization of equation (3), finding G^* needs some comment for those unfamiliar with convex analysis. Using the definition of the Legendre transform for $G(x) = 0$, we have

$$G^*(q) = \max_x \langle q, x\rangle_I. \tag{18}$$

There are two possibilities: (1) $q = \mathbf{0}_I$, in which case the maximum value of $\langle q, x\rangle_I$ is 0, and (2) $q \neq \mathbf{0}_I$, in which case this inner product can increase without bound, resulting in a maximum value of ∞. Putting these two cases together yields the indicator function in

equation (17). With F, G and their conjugates, the optimization problem dual to equation (11) can be written from equation (2):

$$\max_p \left\{-\tfrac{1}{2}\|p\|_2^2 - \langle p, g\rangle_D - \delta_{\mathbf{0}_I}(-A^T p)\right\}. \tag{19}$$

For deriving the CP algorithm instance, it is not strictly necessary to have this dual problem, but it is useful for evaluating convergence.

The CP algorithm solves equations (11) and (19) simultaneously. In principle, the values of the primal and dual objective functions provide a test of convergence. During the iteration, the objective of the primal problem will be greater than the objective of the dual problem, and when the solutions of the respective problems are reached, these objectives will be equal. Comparing the duality gap, i.e. the difference between the primal objective and the dual objective, with 0 thus provides a test of convergence. The presence of the indicator function in the dual problem, however, complicates this test. Due to the negative sign in front of the indicator, when the argument is not the zero vector, this term and therefore the whole dual objective is assigned to a value of $-\infty$. The dual objective achieves a finite, testable value only when the indicator function attains the value of 0, when $A^T p = \mathbf{0}_I$. Effectively, the indicator function becomes a way to write down a constraint in the form of a convex function, in this case an equality constraint. The dual optimization problem can thus alternately be written as a conventional constrained maximization

$$\max_p \left\{-\tfrac{1}{2}\|p\|_2^2 - \langle p, g\rangle_D\right\} \text{ such that } A^T p = \mathbf{0}_I. \tag{20}$$

The convergence check is a bit problematic, because the equality constraint will not likely be strictly satisfied in numerical computation. Instead, we introduce a conditional primal–dual gap (the difference between the primal and dual objectives ignoring the indicator function) given the estimates u' and p':

$$cPD(u', p') = \tfrac{1}{2}\|Au' - g\|_2^2 + \tfrac{1}{2}\|p'\|_2^2 + \langle p', g\rangle_D, \tag{21}$$

and separately monitor $A^T p'$ to see if it is tending to $\mathbf{0}_I$. Note that the conditional primal–dual gap need not be positive, but it should tend to zero.

To finally attain the CP algorithm instance for image reconstruction by least-squares, we derive lines 5 and 6 in algorithm 1. The proximal mapping $prox_\sigma[F^*](y)$, $y \in D$, for this problem results from a quadratic minimization

$$\begin{aligned} prox_\sigma[F^*](y) &= \arg\min_{y'} \left\{\frac{1}{2}\|y'\|_2^2 + \langle y', g\rangle_D + \frac{\|y - y'\|_2^2}{2\sigma}\right\} \\ &= \frac{y - \sigma g}{1 + \sigma}, \end{aligned} \tag{22}$$

and as $G(x) = 0$, $x \in I$, the corresponding proximal mapping is

$$prox_\tau[G](x) = x. \tag{23}$$

Substituting the arguments from the generic algorithm leads to the update steps in listing 2. The constant $L = \|A\|_2$ is the largest singular value of A (see appendix A for details on the power method). Crucial to the implementation of the CP algorithm instance is that A^T be the exact transpose of A, which is a non-trivial matter for tomographic applications, because the projection matrix A is usually computed on-the-fly (Siddon 1985, De Man and Basu 2004, Xu and Mueller 2007). The convergence of the CP algorithm is only guaranteed when A^T is the exact transpose of A, although it may be possible to extend the CP algorithm to mismatched projector/back-projector pairs by employing the analysis in Zeng and Gullberg (2000).

Algorithm 2. Pseudocode for N steps of the least-squares Chambolle–Pock algorithm instance.

$L \leftarrow \|A\|_2;\ \tau \leftarrow 1/L;\ \sigma \leftarrow 1/L;\ \theta \leftarrow 1;\ n \leftarrow 0$
initialize u_0 and p_0 to zero values
$\bar{u}_0 \leftarrow u_0$
repeat
  $p_{n+1} \leftarrow (p_n + \sigma(A\bar{u}_n - g))/(1+\sigma)$
  $u_{n+1} \leftarrow u_n - \tau A^T p_{n+1}$
  $\bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
  $n \leftarrow n+1$
until $n \geqslant N$

This derivation of the CP least-squares algorithm instance illustrates the method on a familiar optimization problem, and it provides a point of comparison with standard algorithms; this quadratic minimization problem can be solved straightforwardly with the basic, linear CG algorithm. Another important point for this particular algorithm instance, where limited projection data can lead to an underdetermined system, is that the CP algorithm will yield a minimizer of the objective $\|Au - g\|_2^2$ which depends on the initial image u_0. In this case, it is recommended to take advantage of the prototyping capability of the CP framework to augment the optimization problem so that it selects a unique image independent of initialization. For example, one often seeks an image closest to either $\mathbf{0}_I$ or a prior image, which can be formulated by adding a quadratic term $\|u\|_2^2$ or $\|u - u_{\text{prior}}\|_2^2$ with a small combination coefficient.

3.1.1. Adding in non-negativity constraints. One of the flexibilities of the CP method becomes apparent in adding bound constraints. While CG is also a flexible tool for dealing with large and small quadratic optimization, modification to include constraints, such as non-negativity, considerably complicates the CG algorithm. For CP, adding in bound constraints is simply a matter of introducing the appropriate indicator function into the primal problem:

$$\min_u \left\{ \tfrac{1}{2}\|Au - g\|_2^2 + \delta_P(u) \right\}, \tag{24}$$

where the set P is all u with non-negative components. Again, we make the mechanical associations with the primal problem (1):

$$F(y) = \tfrac{1}{2}\|y - g\|_2^2, \tag{25}$$

$$G(x) = \delta_P(x), \tag{26}$$

$$x = u, \quad y = Kx, \tag{27}$$

$$K = A. \tag{28}$$

The difference from the unconstrained problem is the function $G(x)$. It turns out that the convex conjugate of $\delta_P(x)$ is

$$\delta_P^*(x) = \delta_P(-x); \tag{29}$$

see appendix B for insight on the convex conjugate of indicator functions. Straight substitution of G^* and F^* into equation (2) yields the dual problem

$$\max_p \left\{ -\tfrac{1}{2}\|p\|_2^2 - \langle p, g\rangle_D - \delta_P(A^T p) \right\}. \tag{30}$$

As a result the conditional primal–dual gap is the same as before. The difference now is that the constraint checks are that $A^T p$ and u should be non-negative.

To derive the algorithm instance, we need the proximal mapping $prox_\tau[G]$, which by definition is

$$prox_\tau[\delta_P](x) = \arg\min_{x'} \left\{ \delta_P(x') + \frac{\|x - x'\|_2^2}{2\tau} \right\}. \tag{31}$$

The indicator in the objective prevents consideration of negative components of x'. The ℓ_2 term can be regarded as a sum over the square difference between components of x and x'; thus, the objective is separable and can be minimized by constructing x' such that $x'_i = x_i$ when $x_i > 0$ and $x'_i = 0$ when $x_i \leqslant 0$. Thus, this proximal mapping becomes a non-negativity thresholding on each component of x:

$$[prox_\tau[\delta_P](x)]_i = [pos(x)]_i \equiv \begin{cases} x_i & x_i > 0 \\ 0 & x_i \leqslant 0 \end{cases}. \tag{32}$$

Substituting into the generic pseudocode yields listing 3. Again, we have $L = \|A\|_2$. The indicator function δ_P leads to the intuitive modification that non-negativity thresholding is introduced in line 6 of listing 3. In this case, the non-negativity constraint in u will be automatically satisfied by all iterates u_n. Upper bound constraints are equally simple to include.

Algorithm 3. Pseudocode for N steps of the least squares with the non-negativity constraint, CP algorithm instance.

1: $L \leftarrow \|A\|_2$; $\tau \leftarrow 1/L$; $\sigma \leftarrow 1/L$; $\theta \leftarrow 1$; $n \leftarrow 0$
2: initialize u_0 and p_0 to zero values
3: $\bar{u}_0 \leftarrow u_0$
4: **repeat**
5: $\quad p_{n+1} \leftarrow (p_n + \sigma(A\bar{u}_n - g))/(1 + \sigma)$
6: $\quad u_{n+1} \leftarrow pos(u_n - \tau A^T p_{n+1})$
7: $\quad \bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
8: $\quad n \leftarrow n + 1$
9: **until** $n \geqslant N$

3.2. Optimization problems based on the TV semi-norm. Optimization problems with the TV semi-norm have received much attention for CT image reconstruction lately because of their potential to provide high quality images from sparse view sampling (Sidky *et al* 2010, 2011, Bian *et al* 2010, Choi *et al* 2010, Ritschl *et al* 2011, Han *et al* 2011, Xia *et al* 2011). The TV semi-norm has been known to be useful for performing edge-preserving regularization, and recent developments in compressive sensing (CS) have sparked even greater interest in the use of this semi-norm. Algorithmwise the TV semi-norm is difficult to handle. Although it is convex, it is not linear, quadratic or even everywhere-differentiable, and the lack of differentiability precludes the use of standard gradient-based optimization algorithms. In this sub-section, we go through, in detail, the derivation of a CP algorithm instance for a TV-regularized least-squares data error norm. We then consider the Kullback–Leibler (KL) data divergence, which is implicitly employed by many iterative algorithms based on maximum likelihood expectation maximization (MLEM). We also consider a data error norm based on ℓ_1 which can have some advantage in reducing the impact of image discretization error, which generally leads to a highly non-uniform error in the data domain. Finally, we derive a CP

algorithm instance for constrained TV-minimization, which is mathematically equivalent to the least-squares-plus-TV problem (Elad 2010), but whose data-error constraint parameter has more physical meaning than the parameter used in the corresponding unconstrained minimization. While the previous CP instances solve optimization problems, which can be solved efficiently by well-known algorithms, the following CP instances are new for the application of CT image reconstruction.

The optimization problem of interest is

$$\min_u \left\{ \tfrac{1}{2} \|Au - g\|_2^2 + \lambda \, \|(|\nabla u|)\|_1 \right\}, \tag{33}$$

where the last term, the ℓ_1-norm of the gradient-magnitude image, is the isotropic TV semi-norm. The spatial-vector image ∇u represents a discrete approximation to the image gradient which is in the vector space V, i.e. the space of spatial-vector-valued image arrays. The expression $|\nabla u|$ is the gradient-magnitude image, an image array whose pixel values are the gradient magnitude at the pixel location. Thus, $\nabla u \in V$ and $|\nabla u| \in I$. Because ∇ is defined in terms of finite differencing, it is a linear transform from an image array to a vector-valued image array, the precise form of which is covered in appendix D. This problem was not explicitly covered in Chambolle and Pock (2011), and we fill in the details here. For this case, matching the primal problem to equation (1) is not as obvious as the previous examples. We recognize in equation (33) that both terms involve a linear transform; thus, the whole objective function can be written in the form $F(Kx)$ with the following assignments:

$$F(y, z) = F_1(y) + F_2(z), \quad F_1(y) = \frac{1}{2}\|y - g\|_2^2, \quad F_2(z) = \lambda \, \|(|z|)\|_1 , \tag{34}$$

$$G(x) = 0, \tag{35}$$

$$x = u, \quad y = Au, \quad z = \nabla u, \tag{36}$$

$$K = \begin{pmatrix} A \\ \nabla \end{pmatrix}, \tag{37}$$

where $u \in I$, $y \in D$ and $z \in V$. Note that $F(y, z)$ is convex because it is the sum of two convex functions. Also the linear transform K takes an image vector x and gives a data vector y and an image gradient vector z. The transpose of K, $K^T = (A^T, -\text{div})$, will produce an image vector from a data vector y and an image gradient vector z:

$$x \leftarrow A^T y - \text{div}\, z, \tag{38}$$

where we use the same convention as in Chambolle and Pock (2011) that $-\text{div} \equiv \nabla^T$; see appendix D.

In order to get the convex conjugate of F we need F_2^*. For readers unfamiliar with the Legendre transform of indicator functions, appendix B illustrates the transform of some common cases. By definition,

$$F_2^*(q) = \max_z \left\{ \langle q, z \rangle_V - \lambda \, \|(|z|)\|_1 \right\}, \tag{39}$$

where $q \in V$, like z, is a vector-valued image array. There are two cases to be considered: (1) the magnitude image $|q|$ at all pixels is less than or equal to λ, i.e. $|q| \in Box(\lambda)$ and (2) the magnitude image $|q|$ has at least one pixel greater than λ, i.e. $|q| \notin Box(\lambda)$. It turns out that for the former case, the maximization in equation (39) yields 0, while the latter case yields ∞. Putting these two cases together, we have

$$F_2^*(q) = \delta_{\text{Box}(\lambda)}(|q|). \tag{40}$$

The conjugates of F and G are

$$F^*(p, q) = \tfrac{1}{2}\|p\|_2^2 + \langle p, g\rangle_D + \delta_{\mathrm{Box}(\lambda)}(|q|), \tag{41}$$

$$G^*(r) = \delta_{\mathbf{0}_I}(r), \tag{42}$$

where $p \in D$, $q \in V$ and $r \in I$.

The problem dual to equation (33) becomes

$$\max_{p,q} \left\{ -\tfrac{1}{2}\|p\|_2^2 - \langle p, g\rangle_D - \delta_{\mathrm{Box}(\lambda)}(|q|) - \delta_{\mathbf{0}_I}(\mathrm{div}\, q - A^T p) \right\}. \tag{43}$$

The resulting conditional primal–dual gap is

$$cPD(u', p', q') = \tfrac{1}{2}\|Au' - g\|_2^2 + \lambda\|(|\nabla u'|)\|_1 + \tfrac{1}{2}\|p'\|_2^2 + \langle p', g\rangle_P \tag{44}$$

with additional constraints $|q'| \in \mathrm{Box}(\lambda)$ and $A^T p' - \mathrm{div}\, q' = \mathbf{0}_I$. The final piece needed for putting together the CP algorithm instance for equation (33) is the proximal mapping

$$prox_\sigma[F^*](y, z) = \left(\frac{y - \sigma g}{1 + \sigma}, \frac{\lambda z}{\max(\lambda \mathbf{1}_I, |z|)} \right). \tag{45}$$

The proximal mapping of the data term was covered previously, and that of the TV term is explained in appendix C. With the necessary pieces in place, the CP algorithm instance for the ℓ_2^2-TV objective can be written in listing 4. Line 6 and the corresponding expression in equation (45) require some explanation, because the division operation is non-standard as the numerator is in V and the denominator is in I. The effect of this line is to threshold the magnitude of the spatial vectors at each pixel in $q_n + \sigma\nabla\bar{u}_n$ to the value λ: spatial vectors larger than λ have their magnitude rescaled to λ. The resulting thresholded spatial-vector image is then assigned to q_{n+1}. Recall that $\mathbf{1}_I$ in line 6 is an image with all pixels set to 1. The operator $|\cdot|$ in this line converts a vector-valued image in V to a magnitude image in I, and the $\max(\lambda\mathbf{1}_I, \cdot)$ operation thresholds the lower bound of the magnitude image to λ pixelwise. Operationally, the division is performed by dividing the spatial vector at each pixel of the numerator by the scalar in the corresponding pixel of the denominator. Another potential source of confusion is computing the magnitude $\|(A, \nabla)\|_2$. The power method for doing this is covered explicitly in appendix A. If it is desired to enforce the positivity constraint, the indicator $\delta_P(u)$ can be added to the primal objective, and the effect of this indicator is the same as for listing 3; namely, the right-hand side of line 7 goes inside the $pos(\cdot)$ operator.

Algorithm 4. Pseudocode for N steps of the ℓ_2^2-TV CP algorithm instance.

1: $L \leftarrow \|(A, \nabla)\|_2$; $\tau \leftarrow 1/L$; $\sigma \leftarrow 1/L$; $\theta \leftarrow 1$; $n \leftarrow 0$
2: initialize u_0, p_0, and q_0 to zero values
3: $\bar{u}_0 \leftarrow u_0$
4: **repeat**
5: $\quad p_{n+1} \leftarrow (p_n + \sigma(A\bar{u}_n - g))/(1 + \sigma)$
6: $\quad q_{n+1} \leftarrow \lambda(q_n + \sigma\nabla\bar{u}_n)/\max(\lambda\mathbf{1}_I, |q_n + \sigma\nabla\bar{u}_n|)$
7: $\quad u_{n+1} \leftarrow u_n - \tau A^T p_{n+1} + \tau div\, q_{n+1}$
8: $\quad \bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
9: $\quad n \leftarrow n + 1$
10: **until** $n \geqslant N$

3.2.1. Alternate data divergences. For a number of reasons motivated by the physical model of imaging systems, it may be of use to formulate optimization problems for CT image reconstruction with alternate data-error terms. A natural extension of the quadratic data divergence is to include a diagonal weighting matrix. The corresponding CP algorithm instance can be easily derived following the steps mentioned above. As pointed out above, the CP method is not limited to quadratic objective functions and other important convex functions can be used. We derive, here, two additional CP algorithm instances. For alternate data divergences we consider the oft-used KL divergence, and one not so commonly used ℓ_1 data-error norm. For the following, we need only analyze the function F_1, as everything else remains the same as for the ℓ_2^2-TV objective in equation (33).

TV plus KL data divergence One data divergence of particular interest for tomographic image reconstruction is KL. Objectives based on KL are what is being optimized in the various forms of MLEM, and it is used often when data noise is a significant physical factor and the data are modeled as being drawn from a multivariate Poisson probability distribution (Barrett and Myers 2004). For the situation where the view sampling is also sparse, it might be of interest to combine a KL data error term with the TV semi-norm in the following primal optimization:

$$\min_u \left\{ \sum_i [Au - g + g \ln g - g \ln(pos(Au))]_i + \delta_P(Au) + \lambda \, \| (|\nabla u|) \|_1 \right\}, \tag{46}$$

where $\sum_i [\cdot]_i$ performs summation over all components of the vector argument. This example proceeds as above except that the F_1 function is different:

$$F_1(y) = \sum_i [y - g + g \ln g - g \ln(pos(y))]_i + \delta_P(y), \tag{47}$$

where $y \in D$ and the function ln operates on the components of its argument. The use of the KL data divergence makes sense only with positive linear systems A and non-negative pixel values u and data g. However, by defining the function over the whole space and using an indicator function to restrict the domain (Rockafellar *et al* 1970), a wide variety of optimization problems can be treated in a uniform manner. Accordingly, δ_P is introduced into the F_1 objective and the *pos* operator is used just so that this objective is defined in the real numbers. The derivation of F_1^*, though mechanical, is a little bit too long to be included here. We simply state the resulting conjugate function:

$$F_1^*(p) = \sum_i [-g \ln pos(\mathbf{1}_D - p)]_i + \delta_P(\mathbf{1}_D - p). \tag{48}$$

The resulting dual problem to equation (46) is thus

$$\max_{p,q} \left\{ \sum_i [g \ln pos(\mathbf{1}_D - p)]_i - \delta_P(\mathbf{1}_D - p) - \delta_{\mathrm{Box}(\lambda)}(|q|) - \delta_{\mathbf{0}_I}(div\, q - A^T p) \right\}. \tag{49}$$

To form the algorithm instance, we need the proximal mapping $prox_\sigma[F_1^*](y)$

$$prox_\sigma[F_1^*](y) = \tfrac{1}{2}(\mathbf{1}_D + y - \sqrt{(y - \mathbf{1}_D)^2 + 4\sigma g}). \tag{50}$$

An interesting point in the derivation, shown partially in appendix C, of $prox_\sigma[F_1^*](y)$ is that the quadratic equation is needed, and the indicator function in $F_1^*(p)$ is used to select the correct (in this case negative) root of the discriminant in the quadratic formula. With the new function F_1, its conjugate and the conjugate's proximal mapping, we can write down the CP algorithm instance. Listing 5 gives the CP algorithm instance minimizing a KL plus TV semi-norm objective.The difference between this algorithm instance and the previous ℓ_2^2-TV

case comes only at the update at line 5. This algorithm instance has the interesting property that the intermediate image estimates u_n can have negative values even though the converged solution will be non-negative. If it is desirable to have the intermediate image estimates be non-negative, the non-negativity constraint can be easily introduced by adding the indicator $\delta_P(u)$ to the primal objective, resulting in the addition of the $pos(\cdot)$ operator at line 7 as was shown in listing 3.

Algorithm 5. Pseudocode for N steps of the KL-TV CP algorithm instance.

1: $L \leftarrow \|(A, \nabla)\|_2;\ \tau \leftarrow 1/L;\ \sigma \leftarrow 1/L;\ \theta \leftarrow 1;\ n \leftarrow 0$
2: initialize u_0, p_0, and q_0 to zero values
3: $\bar{u}_0 \leftarrow u_0$
4: **repeat**
5: $\quad p_{n+1} \leftarrow \frac{1}{2}\left(\mathbf{1}_D + p_n + \sigma A\bar{u}_n - \sqrt{(p_n + \sigma A\bar{u}_n - \mathbf{1}_D)^2 + 4\sigma g}\right)$
6: $\quad q_{n+1} \leftarrow \lambda(q_n + \sigma\nabla\bar{u}_n)/\max(\lambda\mathbf{1}_I, |q_n + \sigma\nabla\bar{u}_n|)$
7: $\quad u_{n+1} \leftarrow u_n - \tau A^T p_{n+1} + \tau div\, q_{n+1}$
8: $\quad \bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
9: $\quad n \leftarrow n + 1$
10: **until** $n \geqslant N$

TV plus ℓ_1 data-error norm The combination of TV semi-norm regularization and ℓ_1 data-error norm has been proposed for image denoising and it has some interesting properties for that purpose (Chan and Esedoglu 2005). This objective is also presented in Chambolle and Pock (2011). For tomography, this combination may be of interest because the ℓ_1 data-error term is an example of a robust fit to the data. The idea of robust approximation is to weakly penalize data that are outliers (Boyd and Vandenberghe 2004). Fitting with the commonly used quadratic error function clearly puts heavy weight on outlying measurements which in some situations can lead to streak artifacts in the images. In particular, for tomographic image reconstruction with a pixel basis, discretization error and metal objects can lead to highly non-uniform error in the data model. The use of the ℓ_1 data-error term may allow for large errors for measurements along the tangent rays to internal structures, where discretization can have a large effect. The ℓ_1 data-error term also puts greater weight on driving small data errors towards zero. The primal problem of interest is

$$\min_u \{\|Au - g\|_1 + \lambda\,\|(|\nabla u|)\|_1\}. \tag{51}$$

For this objective, the function F_1 is

$$F_1(y) = \|y - g\|_1. \tag{52}$$

Computing the convex conjugate F_1^* yields

$$F_1^*(p) = \delta_{\mathrm{Box}(1)}(p) + \langle p, g\rangle_D, \tag{53}$$

and the resulting dual problem is

$$\max_{p,q}\{-\delta_{\mathrm{Box}(1)}(p) - \langle p, g\rangle_D - \delta_{\mathrm{Box}(\lambda)}(|q|) - \delta_{\mathbf{0}_I}(\mathrm{div}\, q - A^T p)\}. \tag{54}$$

The proximal mapping necessary for completing the algorithm instance is

$$prox_\sigma[F_1^*](y) = \frac{y - g\sigma}{\max(\mathbf{1}_D, |y - g\sigma|)}, \tag{55}$$

where $\mathbf{1}_D$ is a data array with each component set to 1 and the max operation is performed componentwise. The corresponding pseudo-code for minimizing equation (51) is given in listing 6, where the only difference between this code and the previous two occurs at line 5. The ability to deal with non-smooth objectives uncomplicates this particular problem substantially. If smoothness were required, there would have to be smoothing parameters on both the ℓ_1 and TV terms, adding two more parameters than necessary to a study of the image properties as a function of the optimization-problem parameters.

Algorithm 6. Pseudocode for N steps of the ℓ_1-TV CP algorithm instance.

1: $L \leftarrow \|(A, \nabla)\|_2$; $\tau \leftarrow 1/L$; $\sigma \leftarrow 1/L$; $\theta \leftarrow 1$; $n \leftarrow 0$
2: initialize u_0, p_0, and q_0 to zero values
3: $\bar{u}_0 \leftarrow u_0$
4: **repeat**
5: $\quad p_{n+1} \leftarrow (p_n + \sigma(A\bar{u}_n - g))/\max(\mathbf{1}_D, |p_n + \sigma(A\bar{u}_n - g)|)$
6: $\quad q_{n+1} \leftarrow \lambda(q_n + \sigma\nabla\bar{u}_n)/\max(\lambda\mathbf{1}_I, |q_n + \sigma\nabla\bar{u}_n|)$
7: $\quad u_{n+1} \leftarrow u_n - \tau A^T p_{n+1} + \tau div\, q_{n+1}$
8: $\quad \bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
9: $\quad n \leftarrow n + 1$
10: **until** $n \geqslant N$

3.2.2. Constrained, TV-minimization. The previous three optimization problems combine a data fidelity term with a TV penalty, and the balance of the two terms is controlled by the parameter λ. An inconvenience of such optimization problems is that it is difficult to physically interpret λ. Focusing on combining an ℓ_2 data-error norm with TV, reformulating equation (33) as a constrained, TV-minimization leads to the following primal problem:

$$\min_u\{\|(|\nabla u|)\|_1 + \delta_{\text{Ball}(\epsilon)}(Au - g)\}, \tag{56}$$

where $\delta_{\text{Ball}(\epsilon)}(Au - g)$ is zero for $\|Au - g\|_2 \leqslant \epsilon$. When $\epsilon > 0$, this problem is equivalent to the unconstrained optimization (33), see e.g. Elad (2010), in the sense that for each positive ϵ there is a corresponding λ yielding the same solution. For this constrained, TV-minimization, the function F_1 is

$$F_1(y) = \delta_{\text{Ball}(\epsilon)}(y - g). \tag{57}$$

The corresponding conjugate is

$$F_1^*(p) = \epsilon\|p\|_2 + \langle p, g\rangle_D, \tag{58}$$

leading to the dual problem

$$\max_{p,q}\{-\epsilon\|p\|_2 - \langle p, g\rangle_D - \delta_{\text{Box}(1)}(|q|) - \delta_{\mathbf{0}_I}(\text{div}\, q - A^T p)\}. \tag{59}$$

Again for the algorithm instance we need the proximal mapping $prox_\sigma[F_1^*]$:

$$prox_\sigma[F_1^*](y) = \max(\|y - \sigma g\|_2 - \sigma\epsilon, 0)\,(y - \sigma g). \tag{60}$$

The main points in deriving this proximal mapping are discussed in appendix C, and it is an example where geometric/symmetry arguments play a large role. Listing 7 shows the algorithm instance solving equation (56), where once again only line 5 is modified.

This algorithm instance essentially achieves the same goal as listing 4; the only difference is that the parameter ϵ has an actual physical interpretation, being the data-error bound.

Algorithm 7. Pseudocode for N steps of the ℓ_2-constrained, TV-minimization CP algorithm instance.

$L \leftarrow \|(A, \nabla)\|_2;\ \tau \leftarrow 1/L;\ \sigma \leftarrow 1/L;\ \theta \leftarrow 1;\ n \leftarrow 0$
initialize u_0, p_0, and q_0 to zero values
$\bar{u}_0 \leftarrow u_0$
repeat
  $p_{n+1} \leftarrow \max(\|p_n + \sigma(A\bar{u}_n - g)\|_2 - \sigma\epsilon, 0)\,(p_n + \sigma(A\bar{u}_n - g))$
  $q_{n+1} \leftarrow (q_n + \sigma\nabla\bar{u}_n)/\max(\mathbf{1}_I, |q_n + \sigma\nabla\bar{u}_n|)$
  $u_{n+1} \leftarrow u_n - \tau A^T p_{n+1} + \tau \operatorname{div} q_{n+1}$
  $\bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
  $n \leftarrow n + 1$
until $n \geqslant N$

4. Demonstration of CP algorithm instances for tomographic image reconstruction

In the previous section, we have derived CP algorithm instances covering many optimization problems of interest to CT image reconstruction. Not only are there the seven optimization problems, but within each case the system model/matrix A, the data g and optimization problem parameters can vary. For each of these, practically infinite number of optimization problems, the corresponding CP algorithm instances are guaranteed to converge (Chambolle and Pock 2011). The purpose of this results section is not to advocate one optimization problem over another, but rather to demonstrate the utility of the CP algorithm for optimization problem prototyping. For this purpose, we present example image reconstructions that could be performed in a study for investigating the impact of matching the data divergence with the data noise model for image reconstruction in breast CT.

4.1. Experiments for sparse-view image reconstruction from simulated CT data

We briefly describe the significance of the experiments, but we point out that the main goal here is to demonstrate the CP algorithm instances. Much of the recent interest in employing the TV semi-norm in optimization problems for CT image reconstruction has been generated by CS. CS seeks to relate sampling conditions on a sensing device with sparsity in the object being scanned. So far, mathematical results have been limited to various types of random sampling (Candès and Wakin 2008). System matrices such as those representing CT projection fall outside of the scope of mathematical results for CS (Sidky *et al* 2010). As a result, the only current option for investigating CS in CT is through numerical experiments with computer phantoms.

A next logical step for bridging theoretical results for CS to actual application is to consider physical factors in the data model. One such factor is a noise model, which can be quite important for low-dose CT applications such as breast CT. While much work has been performed on iterative image reconstruction with various noise models under conditions of full sampling, little is known about the impact of noise on sparse-view image reconstruction. In the following limited study, we set up a breast CT simulation to investigate the impact

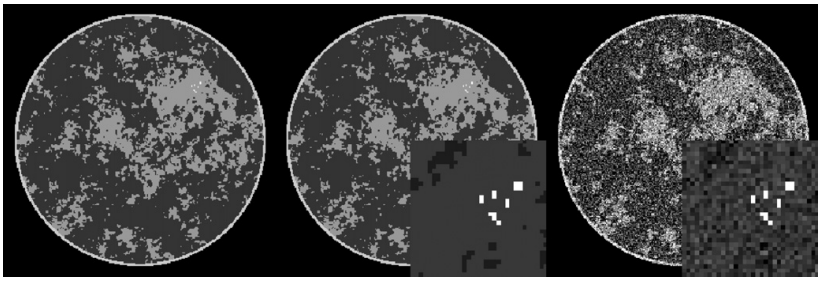


Figure 1. Breast phantom for CT and FBP reconstructed image for a 512-view dataset with Poisson distributed noise. Left: the phantom in the gray scale window [0.95,1.15]; middle: the same phantom with a blow-up on the micro-calcification ROI displayed in the gray scale window [0.9,1.8]; right: the FBP image reconstructed from the noisy data. The middle panel is the reference for all image reconstruction algorithm results. The FBP image is shown only to provide a sense of the noise level.

of correct modeling of data noise with the purpose of demonstrating that the CP algorithm instances can be applied to the CT system.

4.2. Sparse-view reconstruction with a Poisson noise model

For the following study, we employ a digital 256×256 breast phantom, described in Jørgensen *et al* (2011b), Reiser and Nishikawa (2010), and used in our previous study on investigating sufficient sampling conditions for TV-based CT image reconstruction (Jørgensen *et al* 2011a). The phantom models four tissue types: the background fat tissue is assigned a value of 1.0, the modeled fibro-glandular tissue takes a value of 1.1, the outer skin layer is set to 1.15, and the micro-calcifications are assigned values in the range [1.8,2.3].

For the present case, we focus on circular, fan-beam scanning with 60 projections equally distributed over a full 360° angular range. The simulated radius of the x-ray source trajectory is 40 cm with a source–detector distance of 80 cm. The detector sampling consists of 512 bins of size 200 μm. The system matrix for the x-ray projection is computed by the line-intersection method where the matrix elements of A are determined by the length of traversal in each image pixel of each source/detector bin ray. For this phantom under ideal conditions, we have found that an accurate recovery is possible with constrained, TV-minimization with as few as 50 projections. In this study, we add Poisson noise to the data model at a level consistent with what might be expect in a typical breast CT scan. The Poisson noise model is chosen in order to investigate the impact of matching the data-error term to the noise model. For reference, the phantom is shown in figure 1. To have a sense of the noise level, a standard fan-beam filtered back-projection reconstruction is shown alongside the phantom for simulated Poisson noise.

For this noise model, the maximum likelihood method prescribes minimizing the KL data divergence between the available and estimated data. To gauge the importance of selecting a maximum likelihood image, we compare the results from two optimization problems: a KL data divergence plus a TV-penalty, equation (46) above; and a least-squares data error norm plus a TV-penalty, equation (33) above. With the CP framework, these two optimization problems can be easily prototyped: the solutions to both problems can be obtained without worrying about smoothing the TV semi-norm, setting algorithm parameters or proving convergence.

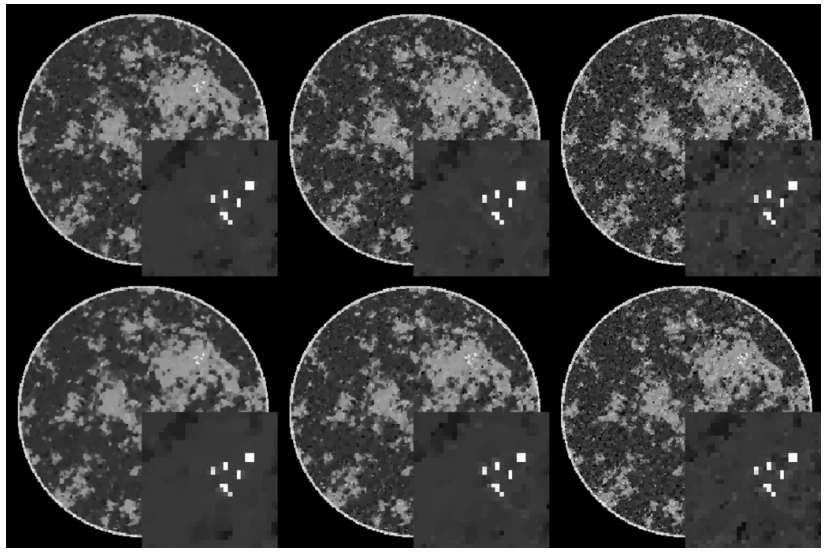


Figure 2. Images reconstructed from 60-view projection data with a Poisson-distributed noise model. The top row of images result from minimizing the ℓ_2^2-TV objective in equation (33) for $\lambda = 1 \times 10^{-4}$, 5×10^{-5} and 2×10^{-5}, going from left to right. The bottom row of images result from minimizing the KL-TV objective in equation (46) for the same values of λ. Note that λ does not necessarily have the same impact on each of these optimization problems. Nevertheless, we see similar trends for the chosen values of λ.

For the phantom and data conditions, described above, the images for different values of the TV-penalty parameter λ are shown in figure 2. An ROI of the micro-calcification cluster is also shown. The overall and ROI images give an impression of two different visual tasks important for breast imaging: discerning the fibro-glandular tissue morphology and detection/classification of micro-calcifications. The images show some difference between the two optimization problems; most notably there is a perceptible reduction in noise in the ROIs from the KL-TV images. A firm conclusion, however, awaits a more complete study with multiple noise realizations.

The most critical feature of the CP algorithm that we wish to promote is the rapid prototyping of a convex optimization problem for CT image reconstruction. The above study is aimed at a combination of using a data divergence based on maximum likelihood estimation with a TV penalty, which takes advantage of sparsity in the gradient magnitude of the underlying object. The CP framework facilitates the use of many other convex optimization problems, particularly those based on some form of sparsity, which often entail some form of the non-smooth ℓ_1-norm. For example, in Sidky *et al* (2010), we have found it useful for sparse-view x-ray phase-contrast imaging to perform image reconstruction with a combination of a least-squares data fidelity term, an ℓ_1-penalty promoting object sparseness, and an image TV constraint to further reduce streak artifacts from angular under-sampling. Under the CP framework, prototyping various combinations of these terms as constrained or unconstrained optimization problems becomes possible and the corresponding derivation of CP algorithm

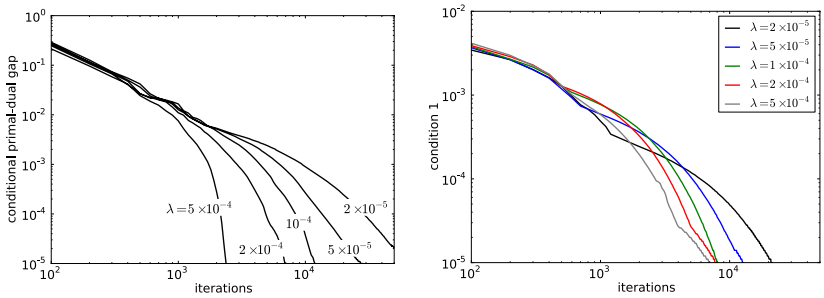


Figure 3. Left: convergence of the conditional primal–dual gap for the CP algorithm instance solving equation (46) for different values of λ. Right: plot indicating agreement with condition 1: $\|div\, q - A^T p\|_\infty$, the magnitude of the largest component of the argument of the last indicator function of equation (49). Collecting all the indicator functions of the primal, equation (46), and dual, equation (49), KL-TV optimization problems, we have four conditions to check in addition to the conditional primal–dual gap: (1) $div\, q - A^T p = \mathbf{0}_I$, (2) $Au \geqslant \mathbf{0}_D$, (3) $p \leqslant \mathbf{1}_D$ and (4) $|q| < \lambda$. The agreement with condition 1 is illustrated in the plot; agreement with condition 2 has a similar dependence; condition 3 is satisfied early on in the iteration; and condition 4 is automatically enforced by the CP algorithm instance for KL-TV. Because the curves are bunched together in the condition 1 plot, they are differentiated in color.

instances follows from the steps described in section 2.2. Alternative, convex data fidelity terms and image constraints motivated by various physical models may also be prototyped.

As a practical matter, though, it is important to have some sense of the convergence of the CP algorithm instances. To this end, we take an in-depth look at individual runs for the KL-TV algorithm instance for CT image reconstruction.

4.3. Iteration dependence of the CP algorithm

Through the methods described above, many useful algorithm instances can be derived for CT image reconstruction. It is obviously important that the resulting algorithm instance reaches the solution of the prescribed optimization problem. To illustrate the convergence of a resulting algorithm instance, we focus on the TV-penalized KL data divergence, equation (46), and plot the conditional primal–dual gap for the different runs with varying λ in figure 3. Included in this figure is a plot indicating the convergence to agreement with the most challenging condition set by the indicator functions in equation (49). For the present results, we terminated the iteration at a conditional primal–dual gap of 10^{-5}, which appears to happen on the scale of thousands of iterations with smaller λ requiring more iterations. Interestingly, a simple pre-conditioned form of the CP algorithm was proposed in Pock and Chambolle (2011), which appears to perform efficiently for small λ. The pre-conditioned CP algorithm instance for this problem is reported in appendix E.

5. Discussion

This paper has presented the application of the CP algorithm to prototyping of optimization problems for CT image reconstruction. The algorithm covers many optimization problems of

interest allowing for non-smooth functions. It also comes with solid convergence criteria to check the image estimates.

The use of the CP algorithm we are promoting here is for prototyping, namely when the image reconstruction algorithm development is at the early stage of determining important factors in formulating the optimization problem. As an example, we illustrated a scenario for sparse-view breast CT considering two different data-error terms. In this stage of development, it is helpful to not have to bother with algorithm parameters, and questions of whether or not the algorithm will converge. After the final optimization problem is determined, the focus shifts from prototyping to efficiency.

Optimization problem prototyping for CT image reconstruction does have its limitations. For example, in the breast CT simulation presented above, a more complete conclusion requires reconstruction from multiple realizations of the data under the Poisson noise model. Additional important dimensions of the study are generation of an ensemble of breast phantoms and considering alternate image representations/projector models. Considering the size of CT image reconstruction systems and huge parameter space of possible optimization problems, it is not yet realistic to completely characterize a particular CT system. But at least we are assured of solving isolated setups and it is conceivable to perform a study along one aspect of the system, i.e. consider multiple realizations of the random data model. Given the current state of affairs for optimization-based image reconstruction, it is crucial that simulations be as realistic as possible. There is great need for realistic phantoms, and data simulation software.

We point out that it is likely at least within the immediate future that optimization-based image reconstruction will have to operate at severely truncated iteration numbers. Current clinical applications of iterative image reconstruction often operate in the range of one to ten iterations, which is likely far too few for claiming that the image estimate is an accurate solution to the designed optimization problem. But at least the ability to prototype an optimization problem can potentially simplify the design phase by separating optimization parameters from algorithm parameters.

Acknowledgments

The authors are grateful to Cyril Riddell and Pierre Vandergheynst for suggesting to look into the Chambolle–Pock algorithm, and to Paul Wolf for checking many of the equations. This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. This work was supported in part by NIH R01 grants CA158446, CA120540 and EB000225. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Appendix A. Computing the norm of K

The matrix norm used for the parameter L in the CP algorithm instances is the largest singular value of K. This singular value can be obtained by the standard power method specified in listing 8. When K represents the discrete x-ray transform, our experience has been that the power method converges to numerical precision in 20 iterations or less. In implementing the CP algorithm instance for TV-penalized minimization, the norm of the combined linear transform $\|(A, \nabla)\|_2$ is needed. For this case, the program is the same as listing 8 where $K^T K x_n$ becomes $A^T A x_n - \text{div}\nabla x_n$; recall that $-\text{div} = \nabla^T$. Furthermore, to obtain s, the explicit computation is $s = \sqrt{\|Ax_{n+1}\|_2^2 + \|\nabla x_{n+1}\|_2^2}$.

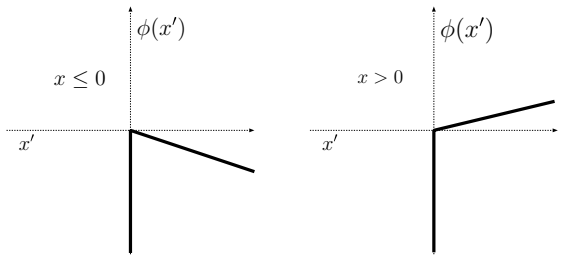


Figure B1. Illustration of the objective function, labeled $\phi(x')$, in the maximization described by equation (B.1). Shown are the two cases discussed in the text.

Algorithm 8. Pseudocode for N steps of the generic power method. The scalar s tends to $\|K\|_2$ as N increases.

initialize $x_0 \in I$ to a non-zero image
$n \leftarrow 0$
repeat
  $x_{n+1} \leftarrow K^T K x_n$
  $x_{n+1} \leftarrow x_{n+1} / \|x_{n+1}\|_2$
  $s \leftarrow \|K x_{n+1}\|_2$
  $n \leftarrow n + 1$
until $n \geqslant N$

Appendix B. The convex conjugate of certain indicator functions of interest illustrated in one dimension

This appendix covers the convex conjugate of a couple of indicator functions in one dimension, serving to illustrate how geometry plays a role in the computation and to provide a mental picture on the conjugate of higher-dimensional indicator functions.

Consider first the indicator $\delta_P(x)$, which is zero for $x \geqslant 0$. The conjugate of this indicator is computed from

$$\delta_P^*(x) = \max_{x'} \phi(x') = \max_{x'} \{x'x - \delta_P(x')\}. \tag{B.1}$$

To perform this maximization, we analyze the cases, $x \leqslant 0$ and $x > 0$, separately. As a visual aid, we plot the objective for these two cases in figure B1. From this figure it is clear that when $x \leqslant 0$, the objective's maximum is attained at $x' = 0$ and this maximum value is 0 (note that this is true even for $x = 0$). When $x > 0$, the objective can increase without bound as x' tends to ∞, resulting in a maximum value of ∞. Putting these two cases together yields

$$\delta_P^*(x) = \delta_P(-x).$$

Generalizing this argument to multi-dimensional x yields equation (29).

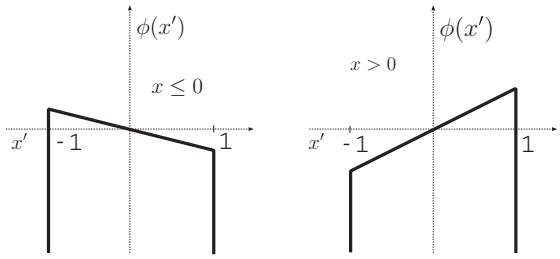


Figure B2. Illustration of the objective function, labeled $\phi(x')$, in the maximization described by equation (B.2). Shown are the two cases discussed in the text.

Next we consider $\delta_{\mathrm{Box}(1)}(x)$, which in one dimension is the same as $\delta_{\mathrm{Ball}(1)}(x)$. This function is zero only for $-1 \leqslant x \leqslant 1$. Its conjugate is computed from

$$\delta^*_{\mathrm{Box}(1)}(x) = \max_{x'} \ \phi(x') = \max_{x'}\{x'x - \delta_{\mathrm{Box}(1)}(x')\}. \tag{B.2}$$

Again we have two cases, $x \leqslant 0$ and $x > 0$, illustrated in figure B2. In the former case, the maximum value of the objective is attained at $x' = -1$, and this maximum value is $-x$. In the latter case, the maximum value is x, and it is attained at $x' = 1$. Hence, we have

$$\delta^*_{\mathrm{Box}(1)}(x) = |x|.$$

For multi-dimensional x, $\delta_{\mathrm{Box}(1)}(x) \neq \delta_{\mathrm{Ball}(1)}(x)$, and this is also reflected in the conjugates

$$\delta^*_{\mathrm{Box}(1)}(x) = \|x\|_1,$$
$$\delta^*_{\mathrm{Ball}(1)}(x) = \|x\|_2.$$

It is also interesting to verify that $\delta^{**}_{\mathrm{Box}(1)}(x)$ is indeed $\delta_{\mathrm{Box}(1)}(x)$ by showing, again in one dimension, that $|x|^* = \delta_{\mathrm{Box}(1)}(x)$. Illustrating this example helps in understanding the convex conjugate of multi-dimensional ℓ_1-based semi-norms. The relevant conjugate is computed from

$$|x|^* = \max_{x'} \phi(x') = \max_{x'}\{x'x - |x'|\}. \tag{B.3}$$

Here, we need to analyze three cases: $x < -1$, $-1 \leqslant x \leqslant 1$ and $x > 1$. The corresponding sketch is in figure B3. The $-|x|$ term in the objective makes an upside-down wedge, and the $x'x$ term serves to tip this wedge. In the second case, the wedge is tipped, but still opens up downward so that the objective is maximized at $x' = 0$, attaining there the value of 0. In the first and third cases, the wedge is tipped so much that part of it points upward and the objective can increase without bound, attaining the value of ∞. Putting these cases together does indeed yield

$$|x|^* = \delta_{\mathrm{Box}(1)}(x).$$

A similar reasoning is used to obtain equation (40) from equation (39).

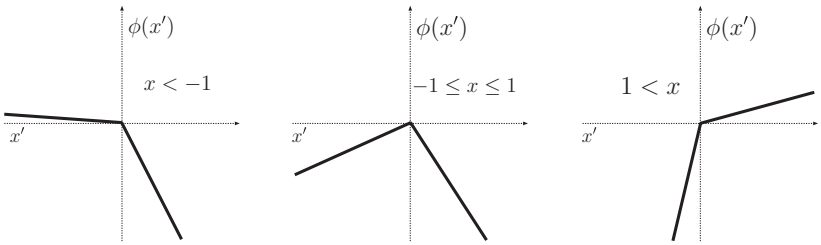


Figure B3. Illustration of the objective function, labeled $\phi(x')$, in the maximization described by equation (B.3). Shown are the three cases discussed in the text.

Appendix C. Computation of important proximal mappings

This appendix fills in important steps in computing some of the proximal mappings in the text, where it is necessary to use geometrical reasoning in addition to setting the gradient of the objective to zero.

The conjugate of the TV semi-norm in equation (40) leads to the following proximal mapping computation:

$$prox_\sigma[F^*](z) = \arg\min_{z'} \left\{ \delta_{\mathrm{Box}(\lambda)}(|z'|) + \frac{\|z - z'\|_2^2}{2\sigma} \right\},$$

where $z, z' \in V$, and the absolute value, $|\cdot|$, of a spatial-vector image V yields an image, in I, of the spatial-vector magnitude. The quadratic term is minimized when $z = z'$, but the indicator function excludes this minimizer when $|z| \notin Box(\lambda)$. To solve this problem, we write the quadratic as a sum over pixels:

$$\frac{\|z - z'\|_2^2}{2\sigma} = \frac{\sum_i |z_i - z_i'|^2}{2\sigma},$$

where i indexes the image pixels and each z_i and z_i' is a spatial-vector. The indicator function places an upper bound on the magnitude of each spatial vector $|z_i'| \leqslant \lambda$. The proximal mapping is built pixel-by-pixel considering two cases: if $|z_i| \leqslant \lambda$, then $prox_\sigma[F^*](z)_i = z_i$; if $|z_i| > \lambda$, then z_i' is chosen to be closest to z_i while respecting $|z_i'| \leqslant \lambda$ which leads to a scaling of the magnitude of z_i and $prox_\sigma[F^*](z)_i = \lambda z_i/|z_i|$. Note that the constant σ does not enter into this calculation. Putting the cases and components all together yields the second part of the proximal mapping in equation (45).

For the KL-TV problem, the proximal mapping for the data term is computed from equation (48):

$$prox_\sigma[F_1^*](p) = \arg\min_{p'} \left\{ \frac{\|p - p'\|_2^2}{2\sigma} - \sum_i [g \ln pos(\mathbf{1}_D - p')]_i + \delta_P(\mathbf{1}_D - p') \right\}.$$

We note that the objective is a smooth function in the positive orthant of $p' \in D$. Accordingly, we differentiate the objective with respect to p' ignoring the $pos(\cdot)$ and indicator functions, keeping in mind that we have to check that the minimizer p' is non-negative. Performing the differentiation and setting to zero yields the following quadratic equation:

$$p'^2 - (\mathbf{1}_D + p)p' + p - \sigma g = 0,$$

and substituting into the quadratic equation yields

$$p' = \tfrac{1}{2}(\mathbf{1}_D + p \pm \sqrt{(\mathbf{1}_D - p)^2 + 4\sigma g}).$$

We have two possible solutions, but it turns out that applying the restriction $\mathbf{1}_D - p' \geqslant 0$ selects the negative root. To see this, we evaluate $\mathbf{1}_D - p'$ at both roots:

$$\mathbf{1}_D - p' = \tfrac{1}{2}(\mathbf{1}_D - p \mp \sqrt{(\mathbf{1}_D - p)^2 + 4\sigma g}).$$

Using the fact that the data are non-negative, we have

$$\sqrt{(\mathbf{1}_D - p)^2 + 4\sigma g} \geqslant |\mathbf{1}_D - p|;$$

the positive root clearly leads to possible negative values for $\mathbf{1}_D - p'$ while the negative root respects $\mathbf{1}_D - p' \geqslant 0$ and yields equation (50).

For the final computation of a proximal mapping, we take a look at the data term of the constrained, TV-minimization problem. From equation (58), the proximal mapping of interest is evaluated by

$$prox_\sigma[F_1^*](p) = \arg\min_{p'} \left\{ \frac{\|p - p'\|_2^2}{2\sigma} + \epsilon \|p'\|_2 + \langle p', g \rangle_D \right\}.$$

Note the first term in the objective is spherically symmetric about p and increasing with distance from p, and the second term is also spherically symmetric about $\mathbf{0}_D$ and increasing with distance from $\mathbf{0}_D$. If just these two terms were present, the minimum would lie on the line segment between $\mathbf{0}_D$ and p. The third term, however, complicates the situation a little. We note that this term is linear in p', and it can be combined with the first term by completing the square. Performing this manipulation and ignoring constant terms (independent of p') yields

$$prox_\sigma[F_1^*](p) = \arg\min_{p'} \left\{ \frac{\|p - p' - \sigma g\|_2^2}{2\sigma} + \epsilon \|p'\|_2 \right\}.$$

By the geometric considerations discussed above, the minimizer lies on the line segment between $\mathbf{0}_D$ and $p - \sigma g$. Analyzing this one-dimensional minimization leads to equation (60).

Appendix D. The finite-differencing form of the image gradient and divergence

In this appendix, we write down the explicit forms of the finite differencing approximations of ∇ and $-\mathrm{div}$ in two dimensions used in this paper. We use $x \in I$ to represent an $M \times M$ image and $x_{i,j}$ to refer to the (i, j)th pixel of x. To specify the linear transform ∇, we introduce the differencing images $\Delta_s x \in I$ and $\Delta_t x \in I$:

$$\Delta_s x_{i,j} = \begin{cases} x_{i+1,j} - x_{i,j} & i < M \\ -x_{i,j} & i = M, \end{cases}$$

$$\Delta_t x_{i,j} = \begin{cases} x_{i,j+1} - x_{i,j} & j < M \\ -x_{i,j} & j = M. \end{cases}$$

Using these definitions, ∇ can be written as

$$\nabla x = \begin{pmatrix} \Delta_s x \\ \Delta_t x \end{pmatrix}.$$

With this form of ∇, its transpose $-$div becomes

$$-\text{div}\begin{pmatrix}\Delta_s x\\ \Delta_t x\end{pmatrix} = \{-(\Delta_s x_{i,j} - \Delta_s x_{i-1,j}) - (\Delta_t x_{i,j} - \Delta_t x_{i,j-1}),\, i \text{ and } j \in [1, M]\},$$

where the elements referred to outside the image border are set to zero: $\Delta_s x_{0,j} = \Delta_s x_{i,0} = \Delta_t x_{0,j} = \Delta_t x_{i,0} = 0$. What the particular form of ∇ is in its discrete form is not that important, but it is critical that the discrete forms of $-div$ and ∇ are the transposes of each other.

Appendix E. Preconditioned CP algorithm demonstrated on the KL-TV optimization problem

Chambolle and Pock followed their article, Chambolle and Pock (2011), with a pre-conditioned version of their algorithm that suits our purpose of optimization problem prototyping while potentially improving algorithm efficiency substantially for the ℓ_2^2-TV and KL-TV optimization problems with small λ. The new algorithm replaces the constants σ and τ with vector quantities that are computed directly from the system matrix K, which yields a vector in space Y from a vector in space X. One form of the suggested, diagonal pre-conditioners uses the following weights:

$$\Sigma = \frac{\mathbf{1}_Y}{|K|\mathbf{1}_X}, \tag{E.1}$$

$$\mathrm{T} = \frac{\mathbf{1}_X}{|K|^T\mathbf{1}_Y}, \tag{E.2}$$

where $\Sigma \in Y$, $\mathrm{T} \in X$, and $|K|$ is the matrix formed by taking the absolute value of each element of K. In order to generate the CP algorithm instance incorporating pre-conditioning, the proximal mapping needs to be modified:

$$prox_\Sigma[F](y) = \arg\min_{y'} \left\{F(y') + \frac{1}{2}(y - y')^T\left(\frac{y - y'}{\Sigma}\right)\right\}. \tag{E.3}$$

The second term in this minimization is still quadratic but no longer spherically symmetric. The difficulty in deriving the pre-conditioned CP algorithm instances is similar to that of the original algorithm. On the one hand there is no need for finding $\|K\|_2$, but on the other hand deriving the proximal mapping may become more involved. For the ℓ_2^2-TV and the KL-TV optimization problems, the proximal mapping is simple to derive and it turns out that the mappings can be arrived at by replacing σ by Σ and τ by T.

The gain in efficiency for small λ comes from being able to absorb this parameter into the TV term and allowing Σ to account for the mismatch between TV and data agreement terms. We modify the definitions of ∇ and $-div$ matrices from appendix D:

$$\nabla_\lambda x = \begin{pmatrix}\lambda\Delta_s x\\ \lambda\Delta_t x\end{pmatrix},$$

and

$$-\text{div}_\lambda\begin{pmatrix}\Delta_s x\\ \Delta_t x\end{pmatrix} = \{-\lambda(\Delta_s x_{i,j} - \Delta_s x_{i-1,j}) - \lambda(\Delta_t x_{i,j} - \Delta_t x_{i,j-1}),\, i \text{ and } j \in [1, M]\},$$

where again the elements referred to outside the image border are set to zero: $\Delta_s x_{0,j} = \Delta_s x_{i,0} = \Delta_t x_{0,j} = \Delta_t x_{i,0} = 0$.

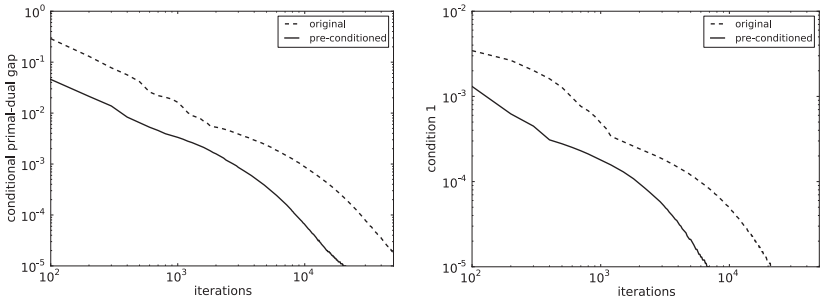


Figure E1. Left: convergence of the partial primal–dual gap for the CP algorithm instance solving equation (46) for $\lambda = 2 \times 10^{-5}$ for the original and pre-conditioned CP algorithm. Right: plot indicating agreement with condition 1 for the KL-TV optimization problem. See figure 3 for explanation.

For a complete example, we write the pre-conditioned CP algorithm instance for KL-TV in listing 9. To illustrate the potential gain in efficiency, we show the condition primal–dual gap as a function of iteration number for the KL-TV problem with $\lambda = 2 \times 10^{-5}$ in figure E1. While we have presented the pre-conditioned CP algorithm as a patch for the small λ case, it really provides an alternative prototyping algorithm and it can be used instead of the original CP algorithm.

Algorithm 9. Pseudocode for N steps of the KL-TV pre-conditioned CP algorithm instance.

$\Sigma_1 \leftarrow \mathbf{1}_D/(|A|\mathbf{1}_I)$; $\Sigma_2 \leftarrow \mathbf{1}_V/(|\nabla_\lambda|\mathbf{1}_I)$; $\mathrm{T} \leftarrow \mathbf{1}_I/(|A^T|\mathbf{1}_D + |div_\lambda|\mathbf{1}_V)$
$\theta \leftarrow 1$; $n \leftarrow 0$
initialize u_0, p_0, and q_0 to zero values
$\bar{u}_0 \leftarrow u_0$
repeat
  $p_{n+1} \leftarrow \frac{1}{2}\left(\mathbf{1}_D + p_n + \Sigma_1 A\bar{u}_n - \sqrt{(p_n + \Sigma_1 A\bar{u}_n - \mathbf{1}_D)^2 + 4\Sigma_1 g}\right)$
  $q_{n+1} \leftarrow (q_n + \Sigma_2 \nabla_\lambda \bar{u}_n)/\max(\mathbf{1}_I, |q_n + \Sigma_2 \nabla_\lambda \bar{u}_n|)$
  $u_{n+1} \leftarrow u_n - \mathrm{T}A^T p_{n+1} + \mathrm{T}div_\lambda\, q_{n+1}$
  $\bar{u}_{n+1} \leftarrow u_{n+1} + \theta(u_{n+1} - u_n)$
  $n \leftarrow n + 1$
until $n \geqslant N$

References

Barrett H H and Myers K J 2004 *Foundations of Image Science* (Hoboken, NJ: Wiley)

Beck A and Teboulle M 2009 Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems *IEEE Trans. Image Process.* **18** 2419–34

Becker S R, Candes E J and Grant M 2011 Templates for convex cone problems with applications to sparse signal recovery *Math. Prog. Comp.* **3** 165–218

Bian J, Siewerdsen J H, Han X, Sidky E Y, Prince J L, Pelizzari C A and Pan X 2010 Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT *Phys. Med. Biol* **55** 6575–99

Boyd S P and Vandenberghe L 2004 *Convex Optimization* (Cambridge: Cambridge University Press)
Candès E J and Wakin M B 2008 An introduction to compressive sampling *IEEE Signal Process. Mag.* **25** 21–30
Chambolle A and Pock T 2011 A first-order primal–dual algorithm for convex problems with applications to imaging *J. Math. Imag. Vis.* **40** 1–26
Chan T F and Esedoglu S 2005 Aspects of total variation regularized L^1 function approximation *SIAM J. Appl. Math.* **65** 1817–37
Chen G H, Tang J and Leng S 2008 Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets *Med. Phys.* **35** 660–3
Choi K, Wang J, Zhu L, Suh T-S, Boyd S and Xing L 2010 Compressed sensing based cone-beam computed tomography reconstruction with a first-order method *Med. Phys.* **37** 5113–25
Combettes P L and Pesquet J C 2008 A proximal decomposition method for solving convex variational inverse problems *Inverse Problems* **24** 065014
De Man B and Basu S 2004 Distance-driven projection and backprojection in three dimensions *Phys. Med. Biol.* **49** 2463–75
Defrise M, Vanhove C and Liu X 2011 An algorithm for total variation regularization in high-dimensional linear problems *Inverse Problems* **27** 065002
Elad M 2010 *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Berlin: Springer)
Erdogan H and Fessler J A 1999 Ordered subsets algorithms for transmission tomography *Phys. Med. Biol* **44** 2835–52
Green P 1984 Iteratively reweighted least squares for maximum likelihood estimation and some robust and resistant alternatives *J. R. Stat. Soc.* B **46** 149–92
Han X, Bian J, Eaker D R, Kline T L, Sidky E Y, Ritman E L and Pan X 2011 Algorithm-enabled low-dose micro-CT imaging *IEEE Trans. Med. Imag.* **30** 606–20
Jensen T L, Jørgensen J H, Hansen P C and Jensen S H 2011 Implementation of an optimal first-order method for strongly convex total variation regularization *BIT* at press (online http://www.springerlink.com/index/10.1007/s10543-011-0359-8)
Jørgensen J H, Sidky E Y and Pan X 2011a Analysis of discrete-to-discrete imaging models for iterative tomographic image reconstruction and compressive sensing *IEEE Trans. Med. Imag.* (arXiv:1109.0629) submitted
Jørgensen J H, Hansen P C, Sidky E Y, Reiser I S and Pan X 2011b Toward optimal x-ray flux utilization in breast CT *Proc. 11th Int. Meeting on Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Potsdam, Germany, 11–15 July 2011)* pp 359–62 (arXiv:1104.1588)
Li M, Yang H and Kudo H 2002 An accurate iterative reconstruction algorithm for sparse objects: application to 3D blood vessel reconstruction from a limited number of projections *Phys. Med. Biol.* **47** 2599–609
McCollough C H, Primak A N, Braun N, Kofler J, Yu L and Christner J 2009 Strategies for reducing radiation dose in CT *Radiol. Clin. N. Am.* **47** 27–40
Nocedal J and Wright S 2006 *Numerical Optimization* 2nd edn (Berlin: Springer)
Pan X, Sidky E Y and Vannier M 2009 Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Problems* **25** 123009
Pock T and Chambolle A 2011 Diagonal preconditioning for first order primal–dual algorithms in convex optimization *IEEE Int. Conf. Computer Vision (ICCV 2011)* pp 1762–69
Ramani S and Fessler J 2011 A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction *IEEE Trans. Med. Imag.* **31** 677–88
Reiser I and Nishikawa R M 2010 Task-based assessment of breast tomosynthesis: effect of acquisition parameters and quantum noise *Med. Phys.* **37** 1591–600
Ritschl L, Bergner F, Fleischmann C and Kachelrieß M 2011 Improved total variation-based CT image reconstruction applied to clinical data *Phys. Med. Biol.* **56** 1545–62
Rockafellar R T 1970 *Convex Analysis* (Princeton, NJ: Princeton University Press)
Siddon R L 1985 Fast calculation of the exact radiological path for a three-dimensional CT array *Med. Phys.* **12** 252–5
Sidky E Y, Anastasio M A and Pan X 2010 Image reconstruction exploiting object sparsity in boundary-enhanced x-ray phase-contrast tomography *Opt. Express* **18** 10404–22
Sidky E Y, Duchin Y, Ullberg C and Pan X 2011 X-ray computed tomography: advances in image formation. A constrained, total-variation minimization algorithm for low-intensity x-ray CT *Med. Phys.* **38** S117–25
Sidky E Y, Kao C-M and Pan X 2006 Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT *J. X-Ray Sci. Tech.* **14** 119–39
Sidky E Y and Pan X 2008 Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization *Phys. Med. Biol.* **53** 4777–807
Xia D, Xiao X, Bian J, Han X, Sidky E Y, Carlo F D and Pan X 2011 Image reconstruction from sparse data in synchrotron-radiation-based microtomography *Rev. Sci. Instrum.* **82** 043706

Xu F and Mueller K 2007 Real-time 3D computed tomographic reconstruction using commodity graphics hardware *Phys. Med. Biol.* **52** 3405–19

Yin W, Osher S, Goldfarb D and Darbon J 2008 Bregman iterative algorithms for $\ell 1$-minimization with applications to compressed sensing *SIAM J. Imag. Sci.* **1** 143–68

Zeng G L and Gullberg G T 2000 Unmatched projector/backprojector pairs in an iterative reconstruction algorithm *IEEE Trans. Med. Imag.* **19** 548–55

Ziegler A, Nielsen T and Grass M 2008 Iterative reconstruction of a region of interest for transmission tomography *Med. Phys.* **35** 1317–27

Appendix F

Implementation of an optimal first-order method for strongly convex total variation regularization

BIT Numer. Math., vol. 52, issue 2, pp. 329–356, 2012.
doi:10.1007/s10543-011-0359-8.
Published online 24 September 2011.

T. L. Jensen, **J. H. Jørgensen**, P. C. Hansen and S. H. Jensen

BIT Numer Math (2012) 52:329–356
DOI 10.1007/s10543-011-0359-8

BIT

Implementation of an optimal first-order method for strongly convex total variation regularization

T.L. Jensen · J.H. Jørgensen · P.C. Hansen · S.H. Jensen

Received: 7 October 2010 / Accepted: 29 August 2011 / Published online: 24 September 2011

Abstract We present a practical implementation of an optimal first-order method, due to Nesterov, for large-scale total variation regularization in tomographic reconstruction, image deblurring, etc. The algorithm applies to μ-strongly convex objective functions with L-Lipschitz continuous gradient. In the framework of Nesterov both μ and L are assumed known—an assumption that is seldom satisfied in practice. We propose to incorporate mechanisms to estimate locally sufficient μ and L during the iterations. The mechanisms also allow for the application to non-strongly convex functions. We discuss the convergence rate and iteration complexity of several first-order methods, including the proposed algorithm, and we use a 3D tomography problem to compare the performance of these methods. In numerical simulations we demonstrate the advantage in terms of faster convergence when estimating the strong convexity parameter μ for solving ill-conditioned problems to high accuracy, in com-

Communicated by Erkki Somersalo.

This work is part of the project CSI: Computational Science in Imaging, supported by grant no. 274-07-0065 from the Danish Research Council for Technology and Production Sciences.

T.L. Jensen · S.H. Jensen
Department of Electronic Systems, Aalborg University, Niels Jernesvej 12, 9220 Aalborg Ø, Denmark

T.L. Jensen
e-mail: tlj@es.aau.dk

S.H. Jensen
e-mail: shj@es.aau.dk

J.H. Jørgensen · P.C. Hansen (✉)
Department of Informatics and Mathematical Modelling, Technical University of Denmark, Building 321, 2800 Lyngby, Denmark
e-mail: pch@imm.dtu.dk

J.H. Jørgensen
e-mail: jakj@imm.dtu.dk

parison with an optimal method for non-strongly convex problems and a first-order method with Barzilai-Borwein step size selection.

Keywords Optimal first-order optimization methods · Strong convexity · Total variation regularization · Tomography

Mathematics Subject Classification (2000) 65K10 · 65R32

1 Introduction

Large-scale discretizations of inverse problems [22] arise in a variety of applications such as medical imaging, non-destructive testing, and geoscience. Due to the inherent instability of these problems, it is necessary to apply regularization in order to compute meaningful reconstructions, and this work focuses on the use of total variation which is a powerful technique when the sought solution is required to have sharp edges (see, e.g., [14, 36] for applications in image reconstruction).

Many total variation algorithms have already been developed, including time marching [36], fixed-point iteration [40], and various minimization-based methods such as sub-gradient methods [1, 15], interior-point methods for second-order cone programming (SOCP) [20], methods exploiting duality [10, 13, 24], and graph-cut methods [11, 18].

The numerical difficulty of a problem depends on the linear forward operator. Most methods are dedicated either to denoising, where the operator is simply the identity, or to deblurring where the operator is represented by a fast transform. For general linear operators with no exploitable matrix structure, such as in tomographic reconstruction, the selection of algorithms is not as large. Furthermore, the systems that arise in real-world tomography applications, especially in 3D, are so large that memory-requirements preclude the use of second-order methods with quadratic convergence.

Recently, Nesterov's optimal first-order method [30, 31] has been adapted to, and analyzed for, a number of imaging problems [16, 41]. In [41] it is shown that Nesterov's method outperforms standard first-order methods by an order of magnitude, but this analysis does not cover tomography problems. A drawback of Nesterov's algorithm (see, e.g., [12]) is the explicit need for the strong convexity parameter and the Lipschitz constant of the objective function, both of which are generally not available in practice.

This paper describes a practical implementation of Nesterov's algorithm, augmented with efficient heuristic methods to estimate the unknown Lipschitz constant and strong convexity parameter. The Lipschitz constant is handled using backtracking, similar to the technique used in [4]. To estimate the unknown strong convexity parameter—which is more difficult—we propose a heuristic based on adjusting an estimate of the strong convexity parameter using a local strong convexity inequality. Furthermore, we equip the heuristic with a restart procedure to ensure convergence in case of an inadequate estimate.

We call the algorithm *UPN* (Unknown Parameter Nesterov) and compare it with two versions of the well-known gradient projection algorithm; *GP*: a simple version

using a backtracking line search for the stepsize and *GPBB*: a more advanced version using Barzilai-Borwein stepsize selection [2] and the nonmonotone backtracking procedure from [21].

We also compare with a variant of the proposed algorithm, UPN_0, where the strong convexity information is not enforced. UPN_0 is optimal among first-order methods for the class of Lipschitz smooth, convex (but *not* strongly convex) functions. There are several other variants of optimal first-order methods for Lipschitz smooth problems, see, e.g., [4, 7, 27, 29–32, 38] and the overview in [6, 38], but they all share similar practical convergence [6, §6.1]. We therefore consider UPN_0 to represent this class of methods. We have implemented the four algorithms in C with a MEX interface to MATLAB, and the software is available from www.imm.dtu.dk/~pch/TVReg/.

Our numerical tests demonstrate that the proposed method *UPN* is significantly faster than *GP*, as fast as *GPBB* for moderately ill-conditioned problems, and significantly faster for ill-conditioned problems. Compared to UPN_0, *UPN* is consistently faster, when solving to high accuracy.

We start with introductions to the discrete total variation problem, to smooth and strongly convex functions, and to some basic first-order methods in Sects. 2, 3, and 4, respectively. Section 5 introduces important inequalities while the new algorithm is described in Sect. 6. Finally, in Sect. 7 we report our numerical experiments with the proposed method applied to an image deblurring problem and a tomographic reconstruction problem.

Throughout the paper we use the following notation. The smallest singular value of a matrix A is denoted $\sigma_{\min}(A)$. The smallest and largest eigenvalues of a symmetric semi-definite matrix M are denoted by $\lambda_{\min}(M)$ and $\lambda_{\max}(M)$. For an optimization problem, f is the objective function, $x^\star$ denotes a minimizer, $f^\star = f(x^\star)$ is the optimum objective, and x is called an ϵ-suboptimal solution if $f(x) - f^\star \leq \epsilon$.

2 The discrete total variation reconstruction problem

The Total Variation (TV) of a real function $\mathcal{X}(t)$ with $t \in \Omega \subset \mathbb{R}^p$ is defined as

$$\mathcal{T}(\mathcal{X}) = \int_\Omega \|\nabla \mathcal{X}(t)\|_2 \, dt. \tag{2.1}$$

Note that the Euclidean norm is not squared, which means that $\mathcal{T}(\mathcal{X})$ is non-differentiable. In order to handle this we consider a smoothed version of the TV functional. Two common choices are to replace the Euclidean norm of the vector z by either $(\|z\|_2^2 + \beta^2)^{1/2}$ or the Huber function

$$\Phi_\tau(z) = \begin{cases} \|z\|_2 - \frac{1}{2}\tau, & \text{if } \|z\|_2 \geq \tau, \\ \frac{1}{2\tau}\|z\|_2^2, & \text{else.} \end{cases} \tag{2.2}$$

In this work we use the latter, which can be considered a prox-function smoothing [31] of the TV functional [5]; thus, the approximated TV functional is given by

$$\mathcal{T}_\tau(\mathcal{X}) = \int_\Omega \Phi_\tau(\nabla \mathcal{X}) \, dt. \tag{2.3}$$

In this work we consider the case $t \in \mathbb{R}^3$. To obtain a discrete version of the TV reconstruction problem, we represent $\mathcal{X}(t)$ by an $N = m \times n \times l$ array X, and we let $x = \text{vec}(X)$. Each element or voxel of the array X, with index j, has an associated matrix (a discrete differential operator) $D_j \in \mathbb{R}^{3\times N}$ such that the vector $D_j\, x \in \mathbb{R}^3$ is the forward difference approximation to the gradient at x_j. By stacking all D_j we obtain the matrix D of dimensions $3N \times N$:

$$D = \begin{pmatrix} D_1 \\ \vdots \\ D_N \end{pmatrix}. \tag{2.4}$$

We use periodic boundary conditions in D, which ensures that only a constant x has a TV of 0. Other choices of boundary conditions could easily be implemented.

When the discrete approximation to the gradient is used and the integration in (2.3) is replaced by summations, the discrete and smoothed TV function is given by

$$T_\tau(x) = \sum_{j=1}^{N} \Phi_\tau(D_j x). \tag{2.5}$$

The gradient $\nabla T_\tau(x) \in \mathbb{R}^N$ of this function is given by

$$\nabla T_\tau(x) = \sum_{j=1}^{N} D_j^T\, D_j\, x / \max\{\tau, \|D_j x\|_2\}. \tag{2.6}$$

We assume that the sought reconstruction has voxel values in the range $[0, 1]$, so we wish to solve a bound-constrained problem, i.e., having the feasible region $\mathcal{Q} = \{x \in \mathbb{R}^N \mid 0 \le x_j \le 1, \forall j\}$. Given a linear system $A\, x \approx b$ where $A \in \mathbb{R}^{M\times N}$ and $N = mnl$, we define the associated *discrete TV regularization problem* as

$$x^\star = \underset{x\in\mathcal{Q}}{\text{argmin}}\ \phi(x), \quad \phi(x) = \frac{1}{2}\|A\, x - b\|_2^2 + \alpha\, T_\tau(x), \tag{2.7}$$

where $\alpha > 0$ is the TV regularization parameter. This is the problem we want to solve, for the case where the linear system of equations arises from discretization of an inverse problem.

3 Smooth and strongly convex functions

To set the stage for the algorithm development in this paper, we consider the convex optimization problem $\min_{x\in\mathcal{Q}} f(x)$ where f is a convex function and $\mathcal{Q}$ is a convex set. We recall that a continuously differentiable function f is convex if

$$f(x) \ge f(y) + \nabla f(y)^T (x - y), \quad \forall x, y \in \mathbb{R}^N. \tag{3.1}$$

Definition 3.1 A continuously differentiable convex function f is said to be *strongly convex* with *strong convexity parameter* μ if there exists a $\mu > 0$ such that

$$f(x) \geq f(y) + \nabla f(y)^T (x - y) + \frac{1}{2}\mu \|x - y\|_2^2, \quad \forall x, y \in \mathbb{R}^N. \tag{3.2}$$

Definition 3.2 A continuously differentiable convex function f has *Lipschitz continuous* gradient with *Lipschitz constant* L, if

$$f(x) \leq f(y) + \nabla f(y)^T (x - y) + \frac{1}{2}L \|x - y\|_2^2, \quad \forall x, y \in \mathbb{R}^N. \tag{3.3}$$

Remark 3.1 The condition (3.3) is equivalent [30, Theorem 2.1.5] to the more standard way of defining Lipschitz continuity of the gradient, namely, through convexity and the condition $\|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x - y\|_2, \forall x, y \in \mathbb{R}^N$.

Remark 3.2 Lipschitz continuity of the gradient is a smoothness requirement on f. A function f that satisfies (3.3) is said to be smooth, and L is also known as the *smoothness constant*.

The set of functions that satisfy (3.2) and (3.3) is denoted $\mathcal{F}_{\mu,L}$. It is clear that $\mu \leq L$ and also that if $\mu_1 \geq \mu_0$ and $L_1 \leq L_0$ then $f \in \mathcal{F}_{\mu_1,L_1} \Rightarrow f \in \mathcal{F}_{\mu_0,L_0}$. Given fixed choices of μ and L, we introduce the ratio $Q = L/\mu$ (sometimes referred to as the "modulus of strong convexity" [28] or the "condition number for f" [30]) which is an upper bound for the condition number of the Hessian matrix. The number Q plays a major role for the convergence rate of the methods we will consider.

Lemma 3.1 *For the quadratic function* $f(x) = \frac{1}{2}\|Ax - b\|_2^2$ *with* $A \in \mathbb{R}^{M\times N}$ *we have*

$$L = \|A\|_2^2, \qquad \mu = \lambda_{\min}\left(A^T A\right) = \begin{cases} \sigma_{\min}(A)^2, & \text{if } \operatorname{rank}(A) = N, \\ 0, & \text{else}, \end{cases} \tag{3.4}$$

and if $\operatorname{rank}(A) = N$ *then* $Q = \kappa(A)^2$, *the square of the condition number of* A.

Proof Follows from $f(x) = f(y) + (Ay - b)^T A(x - y) + \frac{1}{2}(x - y)^T A^T A(x - y)$, the second order Taylor expansion of f about y, where equality holds for quadratic f. □

Lemma 3.2 *For the smoothed TV function* (2.5) *we have*

$$L = \|D\|_2^2/\tau, \qquad \mu = 0, \tag{3.5}$$

where $\|D\|_2^2 \leq 12$ *in the* 3*D case*.

Proof The result for L follows from [31, Theorem 1] since the smoothed TV functional can be written as [5, 16]

$$T_\tau(x) = \max_u \left\{ u^T Dx - \frac{\tau}{2}\|u\|_2^2 : \|u_i\|_2 \leq 1, \ \forall i = 1, \ldots, N \right\}$$

with $u = (u_1^T, \ldots, u_N^T)^T$ stacked according to D. The inequality $\|D\|_2^2 \leq 12$ follows from a straightforward extension of the proof in the Appendix of [16]. For μ pick $y = \alpha e \in \mathbb{R}^N$ and $x = \beta e \in \mathbb{R}^N$, where $e = (1, \ldots, 1)^T$, and $\alpha \neq \beta \in \mathbb{R}$. Then we get $T_\tau(x) = T_\tau(y) = 0$, $\nabla T_\tau(y) = 0$ and obtain

$$\frac{1}{2}\mu\|x - y\|_2^2 \leq T_\tau(x) - T_\tau(y) - \nabla T_\tau(y)^T(x - y) = 0,$$

and hence $\mu = 0$. □

Theorem 3.1 *For the function $\phi(x)$ defined in (2.7) we have a strong convexity parameter $\mu = \lambda_{\min}(A^T A)$ and Lipschitz constant $L = \|A\|_2^2 + \alpha\|D\|_2^2/\tau$. If $\operatorname{rank}(A) < N$ then $\mu = 0$, otherwise $\mu = \sigma_{\min}(A)^2 > 0$ and*

$$Q = \kappa(A)^2 + \frac{\alpha}{\tau}\frac{\|D\|_2^2}{\sigma_{\min}(A)^2}, \tag{3.6}$$

where $\kappa(A) = \|A\|_2/\sigma_{\min}(A)$ is the condition number of A.

Proof Assume $\operatorname{rank}(A) = N$ and consider $f(x) = g(x) + h(x)$ with $g \in \mathcal{F}_{\mu_g, L_g}$ and $h \in \mathcal{F}_{\mu_h, L_h}$. Then $f \in \mathcal{F}_{\mu_f, L_f}$, where $\mu_f = \mu_g + \mu_h$ and $L_f = L_g + L_h$. From μ_f and L_f and using Lemmas 3.1 and 3.2 with $g(x) = \frac{1}{2}\|Ax - b\|_2^2$ and $h(x) = \alpha T_\tau(x)$ we obtain the condition number for ϕ given in (3.6). If $\operatorname{rank}(A) < N$ then the matrix $A^T A$ has at least one zero eigenvalue, and thus $\mu = 0$. □

Remark 3.3 Due to the inequalities used to derive (3.6), there is no guarantee that the given μ and L are the tightest possible for ϕ.

4 Some basic first-order methods

A basic first-order method is the gradient projection method of the form

$$x^{(k+1)} = P_{\mathcal{Q}}\big(x^{(k)} - p_k \nabla f\big(x^{(k)}\big)\big), \quad k = 0, 1, 2, \ldots, \tag{4.1}$$

where $P_{\mathcal{Q}}$ is the Euclidean projection onto the convex set $\mathcal{Q}$ [30]. The following theorem summarizes the convergence properties.

Theorem 4.1 *Let $f \in \mathcal{F}_{\mu,L}$, $p_k = 1/L$ and $x^\star \in \mathcal{Q}$ be the constrained minimizer of f, then for the gradient projection method (4.1) we have*

$$f\big(x^{(k)}\big) - f^\star \leq \frac{L}{2k}\|x^{(0)} - x^\star\|_2^2. \tag{4.2}$$

Moreover, if $\mu \neq 0$ then

$$f\big(x^{(k)}\big) - f^\star \leq \left(1 - \frac{\mu}{L}\right)^k \big(f\big(x^{(0)}\big) - f^\star\big). \tag{4.3}$$

Proof The two bounds follow from [39] and [28, §7.1.4], respectively. □

To improve the convergence of the gradient (projection) method, Barzilai and Borwein [2] suggested a scheme in which the step $p_k \nabla f(x^{(k)})$ provides a simple and computationally cheap approximation to the Newton step $(\nabla^2 f(x^{(k)}))^{-1} \nabla f(x^{(k)})$. For general unconstrained problems with $f \in \mathcal{F}_{\mu,L}$, possibly with $\mu = 0$, non-monotone line search combined with the Barzilai-Borwein (BB) strategy produces algorithms that converge [35]; but it is difficult to give a precise iteration complexity for such algorithms. For strictly quadratic unconstrained problems the BB strategy requires $\mathcal{O}(Q \log \epsilon^{-1})$ iterations to obtain an ϵ-suboptimal solution [17]. In [19] it was argued that, in practice, $\mathcal{O}(Q \log \epsilon^{-1})$ iterations "is the best that could be expected". This comment is also supported by the statement in [30, p. 69] that all "reasonable step-size rules" have the same iteration complexity as the standard gradient method. Note that the classic gradient method (4.1) has $\mathcal{O}(L/\epsilon)$ complexity for $f \in \mathcal{F}_{0,L}$. To summarize, when using the BB strategy we should not expect better complexity than $\mathcal{O}(L/\epsilon)$ for $f \in \mathcal{F}_{0,L}$, and $\mathcal{O}(Q \log \epsilon^{-1})$ for $f \in \mathcal{F}_{\mu,L}$.

In Algorithm 1 we give the (conceptual) algorithm *GPBB*, which implements the BB strategy with non-monotone line search [8, 42] using the backtracking procedure from [21] (initially combined in [35]). The algorithm needs the real parameter $\sigma \in [0, 1]$ and the nonnegative integer K, the latter specifies the number of iterations over which an objective decrease is guaranteed.

An alternative approach is to consider first-order methods with optimal complexity. The optimal complexity is defined as the worst-case complexity for a first-order method applied to any problem in a certain class [28, 30] (there are also more technical aspects involving the problem dimensions and a black-box assumption). In this paper we focus on the classes $\mathcal{F}_{0,L}$ and $\mathcal{F}_{\mu,L}$.

Algorithm 1: *GPBB*

input : $x^{(0)}$, K
output: $x^{(k+1)}$

$p_0 = 1$;
for $k = 0, 1, 2, \ldots$ **do**
  `// BB strategy`
  if $k > 0$ **then**
    $p_k \leftarrow \frac{\|x^{(k)} - x^{(k-1)}\|_2^2}{(x^{(k)} - x^{(k-1)})^T (\nabla f(x^{(k)}) - \nabla f(x^{(k-1)}))}$;
  $\beta \leftarrow 0.95$;
  $\bar{x} \leftarrow P_{\mathcal{Q}}(x^{(k)} - \beta p_k \nabla f(x^{(k)}))$;
  $\hat{f} \leftarrow \max\{f(x^{(k)}), f(x^{(k-1)}), \ldots, f(x^{(k-K)})\}$;
  while $f(\bar{x}) \geq \hat{f} - \sigma \nabla f(x^{(k)})^T (x^{(k)} - \bar{x})$ **do**
    $\beta \leftarrow \beta^2$;
    $\bar{x} \leftarrow P_{\mathcal{Q}}(x^{(k)} - \beta p_k \nabla f(x^{(k)}))$;
  $x^{(k+1)} \leftarrow \bar{x}$;

Recently there has been a great deal of interest in optimal first-order methods for convex optimization problems with $f \in \mathcal{F}_{0,L}$ [3, 38]. For this class it is possible to reach an ϵ-suboptimal solution within $\mathcal{O}(\sqrt{L/\epsilon})$ iterations. Nesterov's methods can be used as stand-alone optimization algorithm, or in a composite objective setup [4, 32, 38], in which case they are called accelerated methods (because the designer violates the black-box assumption). Another option is to apply optimal first-order methods to a smooth approximation of a non-smooth function leading to an algorithm with $\mathcal{O}(1/\epsilon)$ complexity [31]; for practical considerations, see [5, 16].

Optimal methods specific for the function class $\mathcal{F}_{\mu,L}$ with $\mu > 0$ are also known [29, 30]; see also [32] for the composite objective version. However, these methods have gained little practical consideration; for example in [32] all the simulations are conducted with $\mu = 0$. Optimal methods require $\mathcal{O}(\sqrt{Q}\log\epsilon^{-1})$ iterations while the classic gradient method requires $\mathcal{O}(Q\log\epsilon^{-1})$ iterations [28, 30]. For quadratic problems, the conjugate gradient method achieves the same iteration complexity as the optimal first-order method [28].

In Algorithm 2 we state the basic optimal method *Nesterov* [30] with known μ and L; it requires an initial $\theta_0 \geq \sqrt{\mu/L}$. Note that it uses two sequences of vectors, $x^{(k)}$ and $y^{(k)}$. The convergence rate is provided by the following theorem.

Theorem 4.2 *If $f \in \mathcal{F}_{\mu,L}$, $1 > \theta_0 \geq \sqrt{\mu/L}$, and $\gamma_0 = \frac{\theta_0(\theta_0 L - \mu)}{1-\theta_0}$, then for algorithm Nesterov we have*

$$f\left(x^{(k)}\right) - f^\star \leq \frac{4L}{(2\sqrt{L} + k\sqrt{\gamma_0})^2}\left(f\left(x^{(0)}\right) - f^\star + \frac{\gamma_0}{2}\|x^{(0)} - x^\star\|_2^2\right). \tag{4.4}$$

Moreover, if $\mu \neq 0$ then

$$f\left(x^{(k)}\right) - f^\star \leq \left(1 - \sqrt{\frac{\mu}{L}}\right)^k \left(f\left(x^{(0)}\right) - f^\star + \frac{\gamma_0}{2}\|x^{(0)} - x^\star\|_2^2\right). \tag{4.5}$$

Proof See [30, (2.2.19), Theorem 2.2.3] and Appendix A for an alternative proof. □

Except for different constants Theorem 4.2 mimics the result in Theorem 4.1, with the crucial differences that the denominator in (4.4) is squared and μ/L in (4.5) has

Algorithm 2: *Nesterov*

input : $x^{(0)}$, μ, L, θ_0
output: $x^{(k+1)}$

$y^{(0)} \leftarrow x^{(0)}$;
for $k = 0, 1, 2, \ldots$ **do**
  $x^{(k+1)} \leftarrow P_{\mathcal{Q}}(y^{(k)} - L^{-1}\nabla f(y^{(k)}))$;
  $\theta_{k+1} \leftarrow$ positive root of $\theta^2 = (1-\theta)\theta_k^2 + \frac{\mu}{L}\theta$;
  $\beta_k \leftarrow \theta_k(1-\theta_k)/(\theta_k^2 + \theta_{k+1})$;
  $y^{(k+1)} \leftarrow x^{(k+1)} + \beta_k(x^{(k+1)} - x^{(k)})$;

a square root. Comparing the convergence rates in Theorems 4.1 and 4.2, we see that the rates are linear but differ in the linear rate, Q^{-1} and $\sqrt{Q^{-1}}$, respectively. For ill-conditioned problems, it is important whether the complexity is a function of Q or $\sqrt{Q}$, see, e.g., [28, §7.2.8], [7]. This motivates the interest in specialized optimal first-order methods for solving ill-conditioned problems.

5 First-order inequalities for the gradient map

For unconstrained convex problems the (norm of the) gradient is a measure of how close we are to the minimum, through the first-order optimality condition, cf. [9]. For constrained convex problems $\min_{x\in\mathcal{Q}} f(x)$ there is a similar quantity, namely, the *gradient map* defined by

$$G_\nu(x) = \nu(x - P_{\mathcal{Q}}(x - \nu^{-1}\nabla f(x))). \tag{5.1}$$

Here $\nu > 0$ is a parameter and ν^{-1} can be interpreted as the step size of a gradient step. The gradient map is a generalization of the gradient to constrained problems in the sense that if $\mathcal{Q} = \mathbb{R}^N$ then $G_\nu(x) = \nabla f(x)$, and the equality $G_\nu(x^\star) = 0$ is a necessary and sufficient optimality condition [39]. In what follows we review and derive some important first-order inequalities which will be used to analyze the proposed algorithm. We start with a rather technical result.

Lemma 5.1 *Let $f \in \mathcal{F}_{\mu,L}$, fix $x \in \mathcal{Q}$, $y \in \mathbb{R}^N$, and set $x^+ = P_{\mathcal{Q}}(y - \bar{L}^{-1}\nabla f(y))$, where $\bar{\mu}$ and $\bar{L}$ are related to x, y and x^+ by the inequalities*

$$f(x) \geq f(y) + \nabla f(y)^T(x - y) + \frac{1}{2}\bar{\mu}\|x - y\|_2^2, \tag{5.2}$$

$$f(x^+) \leq f(y) + \nabla f(y)^T(x^+ - y) + \frac{1}{2}\bar{L}\|x^+ - y\|_2^2. \tag{5.3}$$

Then

$$f\left(x^+\right) \leq f(x) + G_{\bar{L}}(y)^T(y - x) - \frac{1}{2}\bar{L}^{-1}\|G_{\bar{L}}(y)\|_2^2 - \frac{1}{2}\bar{\mu}\|y - x\|_2^2. \tag{5.4}$$

Proof Follows directly from [30, Theorem 2.2.7]. □

Note that if $f \in \mathcal{F}_{\mu,L}$, then in Lemma 5.1 we can always select $\bar{\mu} = \mu$ and $\bar{L} = L$ to ensure that the inequalities (5.2) and (5.3) are satisfied. However, for specific x, y and x^+, there can exist $\bar{\mu} \geq \mu$ and $\bar{L} \leq L$ such that (5.2) and (5.3) hold. We will use these results to design an algorithm for unknown parameters μ and L.

The lemma can be used to obtain the following lemma. The derivation of the bounds is inspired by similar results for composite objective functions in [32], and the second result is similar to [30, Corollary 2.2.1].

Lemma 5.2 *Let $f \in \mathcal{F}_{\mu,L}$, fix $y \in \mathbb{R}^N$, and set $x^+ = P_{\mathcal{Q}}(y - \bar{L}^{-1}\nabla f(y))$. Let $\bar{\mu}$ and $\bar{L}$ be selected in accordance with (5.2) and (5.3) respectively. Then*

$$\frac{1}{2}\bar{\mu}\|y - x^\star\|_2 \leq \|G_{\bar{L}}(y)\|_2. \tag{5.5}$$

If $y \in \mathcal{Q}$ then

$$\frac{1}{2}\bar{L}^{-1}\|G_{\bar{L}}(y)\|_2^2 \leq f(y) - f\left(x^+\right) \leq f(y) - f^\star. \tag{5.6}$$

Proof From Lemma 5.1 with $x = x^\star$ we use $f(x^+) \geq f^\star$ and obtain

$$\frac{1}{2}\bar{\mu}\|y - x^\star\|_2^2 \leq G_{\bar{L}}(y)^T\left(y - x^\star\right) - \frac{1}{2}\bar{L}^{-1}\|G_{\bar{L}}(y)\|_2^2 \leq \|G_{\bar{L}}(y)\|_2\|y - x^\star\|_2,$$

and (5.5) follows; (5.6) follows from Lemma 5.1 using $y = x$ and $f^\star \leq f(x^+)$. □

As mentioned in the beginning of the section, the results of the corollary say that we can relate the norm of the gradient map at y to the error $\|y - x^*\|_2$ as well as to $f(y) - f^*$. This motivates the use of the gradient map in a stopping criterion:

$$\|G_{\bar{L}}(y)\|_2 \leq \bar{\epsilon}, \tag{5.7}$$

where y is the current iterate, and $\bar{L}$ is linked to this iterate using (5.3). The parameter $\bar{\epsilon}$ is a user-specified tolerance based on the requested accuracy. Lemma 5.2 is also used in the following section to develop a restart criterion to ensure convergence.

6 Nesterov's method with parameter estimation

The parameters μ and L are explicitly needed in *Nesterov*. In case of an unregularized least-squares problem we can in principle compute μ and L as the smallest and largest squared singular value of A, though it might be computational expensive. When a regularization term is present it is unclear whether the tight μ and L can be computed at all. Bounds can be obtained using the result in Theorem 3.1.

A practical approach is to estimate μ and L during the iterations. To this end, we introduce the estimates μ_k and L_k of μ and L in each iteration k. We discuss first how to choose L_k, then μ_k, and finally we state the complete algorithm *UPN* and its convergence properties.

To ensure convergence, the main inequalities (A.6) and (A.7) must be satisfied. Hence, according to Lemma 5.1 we need to choose L_k such that

$$f(x^{(k+1)}) \leq f(y^{(k)}) + \nabla f(y^{(k)})^T(x^{(k+1)} - y^{(k)}) + \frac{1}{2}L_k\|x^{(k+1)} - y^{(k)}\|_2^2. \tag{6.1}$$

This is easily accomplished using *backtracking* on L_k [4]. The scheme, *BT*, takes the form given in Algorithm 3, where $\rho_L > 1$ is an adjustment parameter. If the loop

Algorithm 3: *BT*

input : $y, \bar{L}$
output: $x, \tilde{L}$
$\tilde{L} \leftarrow \bar{L}$;
$x \leftarrow P_Q(y - \tilde{L}^{-1}\nabla f(y))$;
while $f(x) > f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}\tilde{L}\|x-y\|_2^2$ **do**
  $\tilde{L} \leftarrow \rho_L \tilde{L}$;
  $x \leftarrow P_Q(y - \tilde{L}^{-1}\nabla f(y))$;

is executed n_{BT} times, the dominant computational cost of *BT* is $n_{BT} + 2$ function evaluations and 1 gradient evaluation.

For choosing the estimate μ_k we introduce the auxiliary variable $\mu_k^\star$ as the value that causes Definition 3.1 (of strong convexity) for $x^\star$ and $y^{(k)}$ to hold with equality

$$f(x^\star) = f(y^{(k)}) + \nabla f(y^{(k)})^T(x^\star - y^{(k)}) + \frac{1}{2}\mu_k^\star \|x^\star - y^{(k)}\|_2^2. \tag{6.2}$$

From (A.7) with Lemma 5.1 and (A.8) we find that we must choose $\mu_k \leq \mu_k^\star$ to obtain a convergent algorithm. However, as $x^\star$ is, of course, unknown, this task is not straightforward, if at all possible. Instead, we propose a *heuristic* where we select μ_k such that

$$f(x^{(k)}) \geq f(y^{(k)}) + \nabla f(y^{(k)})^T(x^{(k)} - y^{(k)}) + \frac{1}{2}\mu_k \|x^{(k)} - y^{(k)}\|_2^2. \tag{6.3}$$

This is indeed possible since $x^{(k)}$ and $y^{(k)}$ are known iterates. Furthermore, we want the estimate μ_k to be decreasing in order to approach a better estimate of μ. This can be achieved by the choice

$$\mu_k = \min\{\mu_{k-1}, M(x^{(k)}, y^{(k)})\}, \tag{6.4}$$

where we have defined the function

$$M(x,y) = \begin{cases} \frac{f(x) - f(y) - \nabla f(y)^T(x-y)}{\frac{1}{2}\|x-y\|_2^2}, & \text{if } x \neq y, \\ \infty, & \text{else.} \end{cases} \tag{6.5}$$

In words, the heuristic chooses the largest μ_k that satisfies (3.2) for $x^{(k)}$ and $y^{(k)}$, as long as μ_k is not larger than μ_{k-1}. The heuristic is simple and computationally inexpensive and we have found that it is effective for determining a useful estimate. Unfortunately, convergence of *Nesterov* equipped with this heuristic is not guaranteed, since the estimate can be too large. To ensure convergence we include a restart procedure *RUPN* that detects if μ_k is too large, inspired by the approach in [32, §5.3] for composite objectives. *RUPN* is given in Algorithm 4.

Algorithm 4: *RUPN*

$\gamma_1 = \theta_1(\theta_1 L_1 - \mu_1)/(1-\theta_1)$;
if $\mu_k \neq 0$ **and** inequality (6.9) not satisfied **then**
  abort execution of *UPN*;
  restart *UPN* with input $(x^{(k+1)},\ \rho_\mu \mu_k,\ L_k,\ \bar{\epsilon})$;

To analyze the restart strategy, assume that μ_i for all $i = 1, \ldots, k$ are *small enough*, i.e., they satisfy $\mu_i \leq \mu_i^\star$ for $i = 1, \ldots, k$, and μ_k satisfies

$$f(x^\star) \geq f(x^{(0)}) + \nabla f(x^{(0)})^T (x^\star - x^{(0)}) + \frac{1}{2}\mu_k \|x^\star - x^{(0)}\|_2^2. \tag{6.6}$$

When this holds we have the convergence result (using (A.9))

$$f(x^{(k+1)}) - f^\star \leq \prod_{i=1}^{k} (1 - \sqrt{\mu_i / L_i}) \left(f(x^{(1)}) - f^\star + \frac{1}{2}\gamma_1 \|x^{(1)} - x^\star\|_2^2 \right). \tag{6.7}$$

We start from iteration $k = 1$ for reasons which will presented shortly (see Appendix A for details and definitions). If the algorithm uses a projected gradient step from the initial $x^{(0)}$ to obtain $x^{(1)}$, the rightmost factor of (6.7) can be bounded as

$$\begin{aligned}
&f\left(x^{(1)}\right) - f^\star + \frac{1}{2}\gamma_1 \|x^{(1)} - x^\star\|_2^2 \\
&\quad \leq G_{L_0}\left(x^{(0)}\right)^T \left(x^{(0)} - x^\star\right) - \frac{1}{2} L_0^{-1} \left\| G_{L_0}\left(x^{(0)}\right) \right\|_2^2 + \frac{1}{2}\gamma_1 \|x^{(1)} - x^\star\|_2^2 \\
&\quad \leq \left\| G_{L_0}\left(x^{(0)}\right) \right\|_2 \|x^{(0)} - x^\star\|_2 - \frac{1}{2} L_0^{-1} \left\| G_{L_0}\left(x^{(0)}\right) \right\|_2^2 + \frac{1}{2}\gamma_1 \|x^{(0)} - x^\star\|_2^2 \\
&\quad \leq \left(\frac{2}{\mu_k} - \frac{1}{2L_0} + \frac{2\gamma_1}{\mu_k^2} \right) \left\| G_{L_0}\left(x^{(0)}\right) \right\|_2^2.
\end{aligned} \tag{6.8}$$

Here we used Lemma 5.1, and the fact that a projected gradient step reduces the Euclidean distance to the solution [30, Theorem 2.2.8]. Using Lemma 5.2 we arrive at the bound, where $\tilde{L}_{k+1}$ is defined in Algorithm *UPN*:

$$\frac{1}{2}\tilde{L}_{k+1}^{-1} \left\| G_{\tilde{L}_{k+1}}\left(x^{(k+1)}\right) \right\|_2^2 \leq \prod_{i=1}^{k} \left(1 - \sqrt{\frac{\mu_i}{L_i}}\right) \left(\frac{2}{\mu_k} - \frac{1}{2L_0} + \frac{2\gamma_1}{\mu_k^2} \right) \left\| G_{L_0}\left(x^{(0)}\right) \right\|_2^2. \tag{6.9}$$

If the algorithm detects that (6.9) is not satisfied, it can only be because there was at least one μ_i for $i = 1, \ldots, k$ which was *not small enough*. If this is the case, we restart the algorithm with a new $\bar{\mu} \leftarrow \rho_\mu \mu_k$, where $0 < \rho_\mu < 1$ is a parameter, using the current iterate $x^{(k+1)}$ as initial vector.

The complete algorithm *UPN* (Unknown-Parameter Nesterov) is given in Algorithm 5. *UPN* is based on Nesterov's optimal method where we have included backtracking on L_k and the heuristic (6.4). An initial vector $x^{(0)}$ and initial parameters

Algorithm 5: *UPN*

input : $x^{(0)}, \bar{\mu}, \bar{L}, \bar{\epsilon}$
output: $x^{(k+1)}$ or $\tilde{x}^{(k+1)}$

1 $[x^{(1)}, L_0] \leftarrow BT(x^{(0)}, \bar{L})$;
2 $\mu_0 = \bar{\mu},\ y^{(1)} \leftarrow x^{(1)},\ \theta_1 \leftarrow \sqrt{\mu_0/L_0}$;
3 **for** $k = 1, 2, \ldots$ **do**
4 $\quad [x^{(k+1)}, L_k] \leftarrow BT(y^{(k)}, L_{k-1})$;
5 $\quad [\tilde{x}^{(k+1)}, \tilde{L}_{k+1}] \leftarrow BT(x^{(k+1)}, L_k)$;
6 $\quad$ **if** $\|G_{\tilde{L}_{k+1}}(x^{(k+1)})\|_2 \leq \bar{\epsilon}$ **then abort**, **return** $\tilde{x}^{(k+1)}$;
7 $\quad$ **if** $\|G_{L_k}(y^{(k)})\|_2 \leq \bar{\epsilon}$ **then abort**, **return** $x^{(k+1)}$;
8 $\quad \mu_k \leftarrow \min\{\mu_{k-1}, M(x^{(k)}, y^{(k)})\}$;
9 $\quad$ *RUPN*;
10 $\quad \theta_{k+1} \leftarrow$ positive root of $\theta^2 = (1-\theta)\theta_k^2 + (\mu_k/L_k)\,\theta$;
11 $\quad \beta_k \leftarrow \theta_k(1-\theta_k)/(\theta_k^2 + \theta_{k+1})$;
12 $\quad y^{(k+1)} \leftarrow x^{(k+1)} + \beta_k(x^{(k+1)} - x^{(k)})$;

$\bar{\mu} \geq \mu$ and $\bar{L} \leq L$ must be specified along with the requested accuracy $\bar{\epsilon}$. The changes from *Nesterov* to *UPN* are at the following lines:

1: Initial projected gradient step to obtain the bound (6.8) and thereby the bound (6.9) used for the restart criterion.
5: Extra projected gradient step explicitly applied to obtain the stopping criterion $\|G_{\tilde{L}_{k+1}}(x^{(k+1)})\|_2 \leq \bar{\epsilon}$.
6,7: Used to relate the stopping criterion in terms of $\bar{\epsilon}$ to ϵ, see Appendix B.3.
8: The heuristic choice of μ_k in (6.4).
9: The restart procedure for inadequate estimates of μ.

We note that in a practical implementation, the computational work involved in one iteration step of *UPN* may—in the worst case situation—be twice that of one iteration of *GPBB*, due to the two calls to *BT*. However, it may be possible to implement these two calls more efficiently than naively calling *BT* twice. We will instead focus on the iteration complexity of *UPN* given in the following theorem.

Theorem 6.1 *Algorithm UPN, applied to* $f \in \mathcal{F}_{\mu,L}$ *under conditions* $\bar{\mu} \geq \mu$, $\bar{L} \leq L$, $\bar{\epsilon} = \sqrt{(\mu/2)\,\epsilon}$, *stops using the gradient map magnitude measure and returns an* ϵ*-suboptimal solution with iteration complexity*

$$\mathcal{O}\big(\sqrt{Q}\log Q\big) + \mathcal{O}\big(\sqrt{Q}\log \epsilon^{-1}\big). \tag{6.10}$$

Proof See Appendix B. □

The term $\mathcal{O}(\sqrt{Q}\log Q)$ in (6.10) follows from application of several inequalities involving the problem dependent parameters μ and L to obtain the overall bound

(6.9). Algorithm *UPN* is suboptimal since the optimal complexity is $\mathcal{O}(\sqrt{Q}\log\epsilon^{-1})$ but it has the advantage that it can be applied to problems with unknown μ and L.

7 Numerical experiments

7.1 An image deblurring example

We exemplify the use of the algorithm *UPN* to solve a total variation regularized image deblurring problem, where the goal is to determine a sharp image x from a blurred and noisy one $b = Ax + e$. The matrix A models linear motion blur, which renders A sparse, and we use reflexive boundary conditions. For this type of blur no fast transform can be exploited. We add Gaussian noise e with relative noise level $\|e\|_2/\|b\|_2 = 0.01$ and reconstruct using $\alpha = 5.0$ and the default setting of $\tau = 10^{-4} \cdot 255$, where $[0, 255]$ is the dynamic pixel intensity range. The result is shown in Fig. 1. We recognize well-known features of TV-regularized reconstructions: Sharp edges are well-preserved, while fine texture has been over-regularized and has a "patchy" appearance.

To investigate the convergence of the methods, we need the true minimizer $x^\star$ with $\phi(x^\star) = \phi^\star$, which is unknown for the test problem. However, for comparison it is enough to use a reference solution much closer to the true minimizer than the iterates. Thus, to compare the accuracy of the solutions obtained with the accuracy parameter $\bar{\epsilon}$, we use a reference solution computed with accuracy $(\bar{\epsilon} \cdot 10^{-2})$, and with abuse of notation we use $x^\star$ to denote this reference solution.

In Fig. 1 both *UPN* and UPN_0 are seen to be faster than *GP* and *GPBB*, and for a high-accuracy solution *UPN* also outperforms UPN_0. For *UPN*, *GP* and *GPBB* we observe linear rates of convergence, but *UPN* converges much faster. UPN_0 shows a sublinear convergence rate, however the initial phase is steep enough that it takes *UPN* almost 1000 iterations to catch up. We note that the potential of *UPN* seems to be in the case where a high-accuracy solution is needed.

Having demonstrated the performance of the proposed algorithm in an image deblurring problem, we focus in the remainder on a 3D tomography test problem, for which we further study the convergence behavior including the influence of the regularization and smoothing parameters.

7.2 Experiments with 3D tomographic reconstruction

Tomography problems arise in numerous areas, such as medical imaging, nondestructive testing, materials science, and geophysics [23, 26, 33]. These problems amount to reconstructing an object from its projections along a number of specified directions, and these projections are produced by X-rays, seismic waves, or other "rays" penetrating the object in such a way that their intensity is partially absorbed by the object. The absorption thus gives information about the object.

The following generic model accounts for several applications of tomography. We consider an object in 3D with linear attenuation coefficient $\mathcal{X}(t)$, with $t \in \Omega \subset \mathbb{R}^3$.

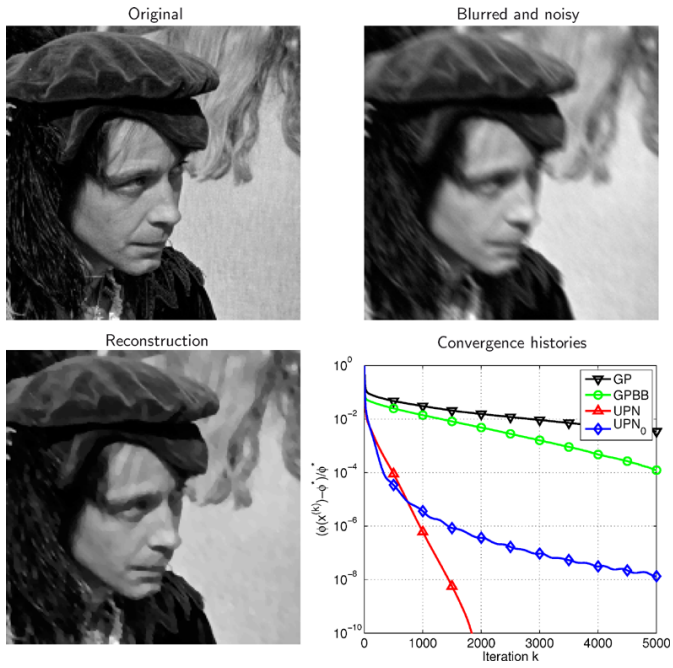


Fig. 1 Example of total variation deblurring for motion blur with reflexive boundary conditions. Methods are Gradient Projection (*GP*), Gradient Projection Barzilai-Borwein (*GPBB*), Unknown Parameter Nesterov (*UPN*), and *UPN* with $\mu_k = 0$ (UPN_0). Both *UPN* and UPN_0 are much faster than *GP* and *GPBB*, and for a high-accuracy solution *UPN* also outperforms UPN_0

The intensity decay b_i of a ray along the line ℓ_i through Ω is governed by a line integral,

$$b_i = \log(I_0/I_i) = \int_{\ell_i} \mathcal{X}(t)\, d\ell = b_i, \tag{7.1}$$

where I_0 and I_i are the intensities of the ray before and after passing through the object. When a large number of these line integrals are recorded, then we are able to reconstruct an approximation of the function $\mathcal{X}(t)$.

We discretize the problem as described in Sect. 2, such that $\mathcal{X}$ is approximated by a piecewise constant function in each voxel in the domain $\Omega = [0, 1] \times [0, 1] \times [0, 1]$. Then the line integral along ℓ_i is computed by summing the contributions from all the voxels penetrated by ℓ_i. If the path length of the ith ray through the jth voxel is denoted by a_{ij}, then we obtain the linear equations

$$\sum_{j=1}^{N} a_{ij} x_j = b_i, \quad i = 1, \ldots, M, \tag{7.2}$$

Fig. 2 *Left*: Two orthogonal slices through the 3D Shepp-Logan phantom discretized on a 43^3 grid used in our test problems. *Middle*: Central horizontal slice. *Right*: Example of solution for $\alpha = 1$ and $\tau = 10^{-4}$. A less smooth solution can be obtained using a smaller α. Original voxel/pixel values are 0.0, 0.2, 0.3 and 1.0. Color range in display is set to [0.1, 0.4] for better contrast

Table 1 Specifications of the two test problems; the object domain consists of $m \times n \times l$ voxels and each projection is a $p \times p$ image. Any zero rows have been purged from A

Problem	$m = n = l$	p	Projections	Dimensions of A	Rank
T1	43	63	37	99361×79507	$=79507$
T2	43	63	13	33937×79507	<79507

where M is the number of rays or measurements and N is the number of voxels. This is a linear system of equations $A\,x = b$ with a sparse coefficient matrix $A \in \mathbb{R}^{M \times N}$.

A widely used test image in medical tomography is the "Shepp-Logan phantom," which consists of a number superimposed ellipses. In the MATLAB function `shepplogan3d` [37] this 2D image is generalized to 3D by superimposing ellipsoids instead. The voxels are in the range [0, 1], and Fig. 2 shows an example with $43 \times 43 \times 43$ voxels.

We construct the matrix A for a parallel-beam geometry with orthogonal projections of the object along directions well distributed over the unit sphere. The projection directions are the direction vectors of so-called *Lebedev quadrature* points on the unit sphere, and the directions are evenly distributed over the sphere; we use the MATLAB implementation `getLebedevSphere` [34]. For setting up the tomography system matrix for a parallel beam geometry, we use the Matlab implementation `tomobox` [25].

This section describes our numerical experiments with the four methods *UPN*, UPN_0, *GP* and *GPBB* applied to the TV regularization problem (2.7). We use the two test problems listed in Table 1, which are representative across a larger class of problems (other directions, number of projections, noise levels, etc.) that we have run simulations with. The smallest eigenvalue of $A^T A$ for *T1* is $2.19 \cdot 10^{-5}$ (as computed by MATLAB's `eigs`), confirming that $\text{rank}(A) = N$ for *T1*. We emphasize that this computation is only conducted to support the analysis of the considered problems since—as we have argued in the introduction—it carries a considerable computational burden to compute. In all simulations we create noisy data from an exact object x_{exact} through the forward mapping $b = Ax_{\text{exact}} + e$, subject to additive Gaussian

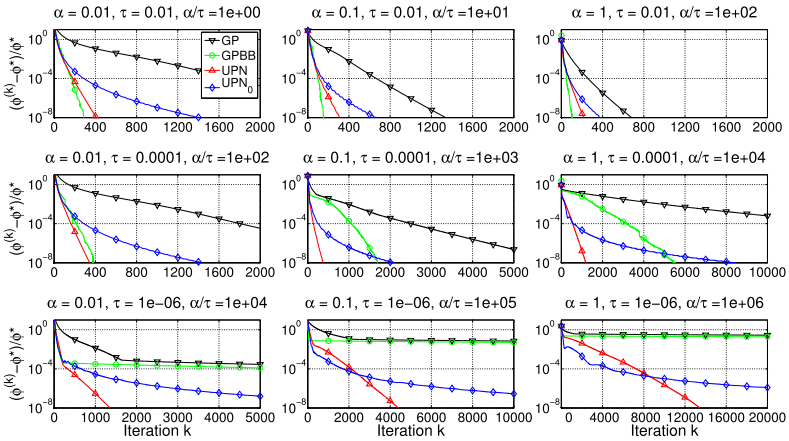


Fig. 3 Convergence histories $(\phi(x^{(k)}) - \phi^\star)/\phi^\star$ vs. k for *T1* with $\alpha = 0.01$, 0.1 and 1 and $\tau = 10^{-2}$, 10^{-4} and 10^{-6}. Methods are Gradient Projection (*GP*), Gradient Projection Barzilai-Borwein (*GPBB*), Unknown Parameter Nesterov (*UPN*), and *UPN* with $\mu_k = 0$ (UPN_0). As the ratio α/τ increases, which implies an increased Q and a computationally more difficult problem, *UPN* and UPN_0 scale significantly better. For high accuracy solutions *UPN* is always competitive

white noise of relative noise level $\|e\|_2/\|b\|_2 = 0.01$. As initial point of the optimization algorithm we use the fifth iteration of the conjugate gradient method applied to the least squares problem.

We compare the algorithm *UPN* with *GP* (the gradient projection method (4.1) with backtracking line search on the step size), *GPBB* and UPN_0. The latter is *UPN* with $\mu_i = 0$ for all $i = 0, \ldots, k$ and $\theta_1 = 1$ and is optimal for the class $\mathcal{F}_{0,L}$.

7.3 Influence of α and τ on the convergence

For a given A the theoretical modulus of strong convexity given in (3.6) varies only with α and τ. We therefore expect better convergence rates (4.3) and (4.5) for smaller α and larger τ. In Fig. 3 we show the convergence histories for *T1* with all combinations of $\alpha = 0.01$, 0.1, 1 and $\tau = 10^{-2}$, 10^{-4}, 10^{-6}.

For low α/τ ratios, i.e., small condition number of the Hessian, *GPBB* and *GP* requires a comparable or smaller number of iterations than *UPN* and UPN_0. As α/τ increases, both *GPBB* and *GP* exhibit slower convergence, while *UPN* is less affected. In all cases *UPN* shows linear convergence, at least in the final stage, while UPN_0 shows sublinear convergence. Due to these observations, we consistently observe that for sufficiently high accuracy, *UPN* requires the lowest number of iterations. This also follows from the theory since *UPN* scales as $\mathcal{O}(\log \epsilon^{-1})$, whereas UPN_0 scales at a higher complexity of $\mathcal{O}(\sqrt{\epsilon^{-1}})$.

We conclude that for small condition numbers there is no gain in using *UPN* compared to *GPBB*. For larger condition numbers, and in particular if a high-accuracy solution is required, *UPN* converges significantly faster. Assume that we were to choose only one of the four algorithms to use for reconstruction across the condition

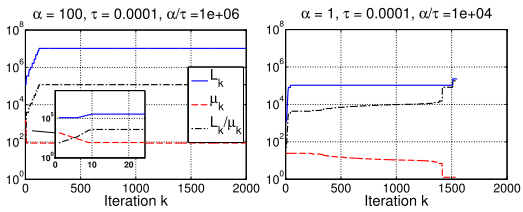


Fig. 4 The μ_k, L_k histories for *T1*. *Left*: $\alpha = 100$ and $\tau = 10^{-4}$. *Right*: $\alpha = 1$ and $\tau = 10^{-4}$

number range. When *UPN* requires the lowest number of iterations, it requires *significantly* fewer, and when not, *UPN* only requires slightly more iterations than the best of the other algorithms. Therefore, *UPN* appears to be the best choice. Obviously, the choice of algorithm also depends on the demanded accuracy of the solution. If only a low accuracy, say $(\phi^{(k)} - \phi^\star)/\phi^\star = 10^{-2}$ is sufficient, all four methods perform more or less equally well.

7.4 Restarts and μ_k and L_k histories

To ensure convergence of *UPN* we introduced the restart functionality *RUPN*. In practice, we almost never observe a restart, e.g., in none of the experiments reported so far a restart occurred. An example where restarts do occur is obtained if we increase α to 100 for *T1* (still $\tau = 10^{-4}$). Restarts occur in the first 8 iterations, and each time μ_k is reduced by a constant factor of $\rho_\mu = 0.7$. In Fig. 4, left, the μ_k and L_k histories are plotted vs. k and the restarts are seen in the zoomed inset as the rapid, constant decrease in μ_k. From the plot we also note that after the decrease in μ_k and an initial increase in L_k, both estimates are constant for the remaining iterations, indicating that the heuristics determine sufficient values.

For comparison the μ_k and L_k histories for *T1* with $\alpha = 1$ and $\tau = 10^{-4}$ are seen in Fig. 4, right. No restarts occurred here, and μ_k decays gradually, except for one final jump, while L_k remains almost constant.

7.5 A non-strongly convex example

Test problem *T2* corresponds to only 13 projections, which causes A to not have full column rank. This leads to $\lambda_{\min}(A^T A) = 0$, and hence $\phi(x)$ is not strongly convex. The optimal convergence rate is therefore given by (4.4); but how does the lack of strong convexity affect *UPN*, which was specifically constructed for strongly convex problems? *UPN* does not recognize that the problem is not strongly convex but simply relies on the heuristic (6.4) at the kth iteration. We investigate the convergence by solving *T2* with $\alpha = 1$ and $\tau = 10^{-4}$. Convergence histories are given in Fig. 5, left. The algorithm *UPN* still converges linearly, although slightly slower than in the *T1* experiment ($\alpha = 1, \tau = 10^{-4}$) in Fig. 3. The algorithms *GP* and *GPBB* converge much more slowly, while at low accuracies UPN_0 is comparable to *UPN*. But the linear convergence makes *UPN* converge faster for high accuracy solutions.

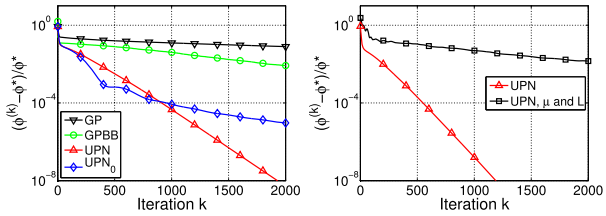


Fig. 5 *Left*: Convergence histories of *GP*, *GPBB*, *UPN* and UPN_0 on *T2* with $\alpha = 1$ and $\tau = 10^{-4}$. *Right*: Convergence histories of *UPN* and *UPN* using true μ and L on *T1* with $\alpha = 1$ and $\tau = 10^{-4}$

7.6 Influence of the heuristic

An obvious question is how the use of the heuristic for estimating μ affects *UPN* compared to *Nesterov*, where μ (and L) are assumed known. From Theorem 3.1 we can compute a strong convexity parameter and a Lipschitz parameter for $\phi(x)$ assuming we know the largest and smallest magnitude eigenvalues of $A^T A$. Recall that these μ and L are not necessarily the tightest possible, according to Remark 3.3. For *T1* we have computed $\lambda_{\max}(A^T A) = 1.52 \cdot 10^3$ and $\lambda_{\min}(A^T A) = 2.19 \cdot 10^{-5}$ (by means of `eigs` in MATLAB). Using $\alpha = 1$, $\tau = 10^{-4}$ and $\|D\|_2^2 \leq 12$ from Lemma 3.2 we fix

$$\mu_k = \lambda_{\min}(A^T A) = 2.19 \cdot 10^{-5}, \qquad L_k = \lambda_{\max}(A^T A) + 12\frac{\alpha}{\tau} = 1.22 \cdot 10^5,$$

for all k, and solve test problem *T1* using *UPN* with the heuristics switched off in favor of these *true* strong convexity and Lipschitz parameters. Convergence histories are plotted in Fig. 5, right.

The convergence is much slower than using *UPN* with the heuristics switched on. We ascribe this behavior to the very large modulus of strong convexity that arise from the true μ and L. It appears that *UPN* works better than the actual degree of strong convexity as measured by μ, by heuristically choosing in each step a μ_k that is sufficient *locally* instead of being restricted to using a *globally* valid μ.

Another question is how much is actually gained in using the heuristic for μ in *UPN* compared to simply using a fixed "guess" throughout the iterations. To answer that question we investigate the number iterations required to obtain $\bar{\epsilon} = 10^{-4}$, 10^{-6} and 10^{-8} solutions for *T1* and *T2* using only the backtracking procedure on L and simply a fixed value $\mu_k \in [10^{-4}, 10^4]$ for all iterations k, see Fig. 6.

The choice of fixed μ_k has a large impact on the required number of iterations, and there is a distinct optimal choice between 1 and 10. Choosing a fixed μ_k away from the optimal one leads to more iterations and the number of additional iterations grows faster for more accurate solutions. For comparison the figure also shows the corresponding number of iterations required by *UPN* plotted as function of the final *UPN*-estimate for μ. For all three *T1* cases *UPN* comes very close to the optimal number of iterations, without demanding an accurate guess of μ by the user. For *T2* we observe similar trends, although *UPN* requires slightly more iterations than with the optimal choice of fixed μ_k.

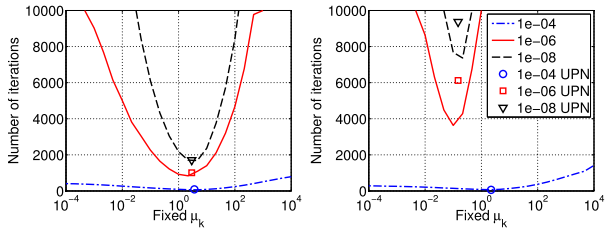


Fig. 6 Number of iterations needed to obtain TV-solutions ($\alpha = 0.01$) to tolerances $\bar{\epsilon} = 10^{-4}$, 10^{-6} and 10^{-8} using fixed μ_k, *left T1*, *right T2*. Also shown are the number iterations needed by *UPN* as function of the final estimate of μ. Choices of μ_k not equal to the unknown optimal value lead to many more iterations. *UPN* needs a near-optimal number of iterations without requiring the user to choose a value for μ

We conclude that there exists a choice of fixed μ_k that gives good performance; however, for an inaccurate guess of this value, the number of iterations will be much higher, in particular if an accurate solution is required. *UPN* avoids the need for such a guess and provides the solution using a near-optimal number of iterations. We emphasize that obtaining a *true* strong convexity parameter μ is not of particular interest here, nor is the final *UPN*-estimate for μ, as the goal is simply to obtain fast convergence.

8 Conclusion

We presented an implementation of an optimal first-order optimization algorithm for large-scale problems, suited for functions that are smooth and strongly convex. While the underlying algorithm by Nesterov depends on knowledge of two parameters that characterize the smoothness and strong convexity, we have implemented methods that estimate these parameters during the iterations, thus making the algorithm of practical use.

We tested the performance of the algorithm and compared it with two variants of the gradient projection algorithm and a variant of the FISTA algorithm. We applied the algorithms to total variation regularized tomographic reconstruction of a generic threedimensional test problem. The tests show that, with regards to the number of iterations, the proposed algorithm is competitive with other first-order algorithms, and superior for difficult problems, i.e., ill-conditioned problems solved to high accuracy. Simulations also show that even for problems that are not strongly convex, in practice we achieve the favorable convergence rate associated with strong convexity. The software is available as a C-implementation with an interface to MATLAB from www.imm.dtu.dk/~pch/TVReg/.

Acknowledgements We wish to thank both referees for their constructive comments which helped improve the presentation of the material.

Appendix A: The optimal convergence rate

Here we provide an analysis of an optimal method for smooth, strongly convex functions without the use of estimation functions as in [30]. This approach is similar to the analysis of optimal methods for smooth functions in [38, 39]. The motivation for the following derivations is to introduce the iteration dependent L_k and μ_k estimates of L and μ. This will support the analysis of how L_k and μ_k should be selected. We start with the following relations to the "hidden" supporting variables $z^{(k)}$ and γ_k [30, pp. 73–75, 89],

$$y^{(k)} - x^{(k)} = \frac{\theta_k \gamma_k}{\gamma_{k+1}}\left(z^{(k)} - y^{(k)}\right), \tag{A.1}$$

$$\begin{aligned} \gamma_{k+1} &= (1-\theta_k)\gamma_k + \theta_k \mu_k = \theta_k^2 L_k, \\ \gamma_{k+1} z^{(k+1)} &= (1-\theta_k)\gamma_k z^{(k)} + \theta_k \mu_k y^{(k)} - \theta_k G_{L_k}\left(y^{(k)}\right). \end{aligned} \tag{A.2}$$

In addition we will make use of the relations

$$\begin{aligned} &\frac{\gamma_{k+1}}{2}\|z^{(k+1)} - y^{(k)}\|_2^2 \\ &\quad = \frac{1}{2\gamma_{k+1}}\Big((1-\theta_k)^2\gamma_k^2\|z^{(k)} - y^{(k)}\|_2^2 \\ &\qquad - 2\theta_k(1-\theta_k)\gamma_k G_{L_k}\left(y^{(k)}\right)^T\left(z^{(k)} - y^{(k)}\right) + \theta_k^2\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2\Big), \end{aligned} \tag{A.3}$$

$$(1-\theta_k)\frac{\gamma_k}{2} - \frac{1}{2\gamma_{k+1}}(1-\theta_k)^2\gamma_k^2 = \frac{(1-\theta_k)\gamma_k\theta_k\mu_k}{2\gamma_{k+1}}. \tag{A.4}$$

which originate from (A.2). We will also later need the relation

$$\begin{aligned} &(1-\theta_k)\frac{\gamma_k}{2}\|z^{(k)} - y^{(k)}\|_2^2 - \frac{\gamma_{k+1}}{2}\|z^{(k+1)} - y^{(k)}\|_2^2 + \theta_k G_{L_k}\left(y^{(k)}\right)^T\left(y^{(k)} - x^\star\right) \\ &\quad = (1-\theta_k)\frac{\gamma_k}{2}\|z^{(k)} - y^{(k)}\|_2^2 - \frac{\gamma_{k+1}}{2}\|z^{(k+1)} - y^{(k)}\|_2^2 \\ &\qquad + \left(-\gamma_{k+1}z^{(k+1)} + (1-\theta_k)\gamma_k z^{(k)} + \theta_k\mu_k y^{(k)}\right)^T\left(y^{(k)} - x^\star\right) \\ &\quad = \left((1-\theta_k)\frac{\gamma_k}{2} - \frac{\gamma_{k+1}}{2} + \theta_k\mu_k\right)\left(y^{(k)}\right)^T y^{(k)} + (1-\theta_k)\frac{\gamma_k}{2}\left(z^{(k)}\right)^T z^{(k)} \\ &\qquad - \frac{\gamma_{k+1}}{2}\left(z^{(k+1)}\right)^T z^{(k+1)} + \gamma_{k+1}\left(z^{(k+1)}\right)^T x^\star \\ &\qquad - (1-\theta_k)\gamma_k\left(z^{(k)}\right)^T x^\star - \theta_k\mu_k\left(y^{(k)}\right)^T x^\star \\ &\quad = (1-\theta_k)\frac{\gamma_k}{2}\left(\|z^{(k)} - x^\star\|_2^2 - \left(x^\star\right)^T x^\star\right) - \frac{\gamma_{k+1}}{2}\left(\|z^{(k+1)} - x^\star\|_2^2 - \left(x^\star\right)^T x^\star\right) \\ &\qquad + \frac{\theta_k\mu_k}{2}\left(\|y^{(k)} - x^\star\|_2^2 - \left(x^\star\right)^T x^\star\right) \\ &\qquad + \left((1-\theta_k)\frac{\gamma_k}{2} - \frac{\gamma_{k+1}}{2} + \frac{\theta_k\mu_k}{2}\right)\left(y^{(k)}\right)^T y^{(k)} \\ &\quad = (1+\theta_k)\frac{\gamma_k}{2}\|z^{(k)} - x^\star\|_2^2 - \frac{\gamma_{k+1}}{2}\|z^{(k)} - x^\star\|_2^2 + \theta_k\frac{\mu_k}{2}\|y^{(k)} - x^\star\|_2^2, \end{aligned} \tag{A.5}$$

where we again used (A.2). We can now start the analysis of the algorithm by considering the inequality in Lemma 5.1,

$$(1-\theta_k) f\left(x^{(k+1)}\right) \leq (1-\theta_k) f\left(x^{(k)}\right) + (1-\theta_k) G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^{(k)}\right) - (1-\theta_k)\frac{1}{2L_k}\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2, \tag{A.6}$$

where we have omitted the strong convexity part, and the inequality

$$\theta_k f\left(x^{(k+1)}\right) \leq \theta_k f\left(x^\star\right) + \theta_k G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^\star\right) - \theta_k \frac{1}{2L_k}\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2 - \theta_k \frac{\mu_k^\star}{2}\|y^{(k)} - x^\star\|_2^2. \tag{A.7}$$

Adding these bounds and continuing, we obtain

$$\begin{aligned}
& f\left(x^{(k+1)}\right) \\
& \leq (1-\theta_k) f\left(x^{(k)}\right) + (1-\theta_k) G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^{(k)}\right) \\
& \quad + \theta_k f^\star + \theta_k G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^\star\right) - \theta_k \frac{\mu_k^\star}{2}\left\|x^\star - y^{(k)}\right\|_2^2 - \frac{1}{2L_k}\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2 \\
& = (1-\theta_k) f\left(x^{(k)}\right) + (1-\theta_k)\frac{\theta_k \gamma_k}{\gamma_{k+1}} G_{L_k}\left(y^{(k)}\right)^T \left(z^{(k)} - y^{(k)}\right) \\
& \quad + \theta_k f^\star + \theta_k G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^\star\right) - \theta_k \frac{\mu_k^\star}{2}\left\|x^\star - y^{(k)}\right\|_2^2 - \frac{1}{2L_k}\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2 \\
& \leq (1-\theta_k) f\left(x^{(k)}\right) + (1-\theta_k)\frac{\theta_k \gamma_k}{\gamma_{k+1}} G_{L_k}\left(y^{(k)}\right)^T \left(z^{(k)} - y^{(k)}\right) \\
& \quad + \theta_k f^\star + \theta_k G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^\star\right) - \theta_k \frac{\mu_k^\star}{2}\left\|x^\star - y^{(k)}\right\|_2^2 - \frac{1}{2L_k}\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2 \\
& \quad + \frac{(1-\theta_k)\theta_k \gamma_k \mu_k}{2\gamma_{k+1}}\left\|z^{(k)} - y^{(k)}\right\|_2^2 \\
& = (1-\theta_k) f\left(x^{(k)}\right) + (1-\theta_k)\frac{\theta_k \gamma_k}{\gamma_{k+1}} G_{L_k}\left(y^{(k)}\right)^T \left(z^{(k)} - y^{(k)}\right) \\
& \quad + \theta_k f^\star + \theta_k G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^\star\right) - \theta_k \frac{\mu_k^\star}{2}\left\|x^\star - y^{(k)}\right\|_2^2 - \frac{1}{2L_k}\left\|G_{L_k}\left(y^{(k)}\right)\right\|_2^2 \\
& \quad + \left((1-\theta_k)\frac{\gamma_k}{2} - \frac{1}{2\gamma_{k+1}}(1-\theta_k)^2 \gamma_k^2\right)\left\|z^{(k)} - y^{(k)}\right\|_2^2 \\
& = (1-\theta_k) f\left(x^{(k)}\right) + (1-\theta_k)\frac{\gamma_k}{2}\left\|z^{(k)} - y^{(k)}\right\|_2^2 - \frac{\gamma_{k+1}}{2}\left\|z^{(k+1)} - y^{(k)}\right\|_2^2 \\
& \quad + \theta_k f^\star + \theta_k G_{L_k}\left(y^{(k)}\right)^T \left(y^{(k)} - x^\star\right) - \theta_k \frac{\mu_k^\star}{2}\left\|x^\star - y^{(k)}\right\|_2^2
\end{aligned}$$

$$= (1-\theta_k) f\left(x^{(k)}\right) + \theta_k f^\star - \theta_k \frac{\mu_k^\star}{2}\left\|x^\star - y^{(k)}\right\|_2^2$$
$$+ (1-\theta_k)\frac{\gamma_k}{2}\|z^{(k)} - x^\star\|_2^2 - \frac{\gamma_{k+1}}{2}\left\|z^{(k+1)} - x^\star\right\|_2^2 + \theta_k \frac{\mu_k}{2}\left\|y^{(k)} - x^\star\right\|_2^2,$$

where we have used (A.1), a trivial inequality, (A.4) , (A.3), (A.2), and (A.5). If $\mu_k \le \mu_k^\star$ then

$$f\left(x^{(k+1)}\right) - f^\star + \frac{\gamma_{k+1}}{2}\left\|z^{(k+1)} - x^\star\right\|_2^2 \le (1-\theta_k)\left(f\left(x^{(k)}\right) - f^\star + \frac{\gamma_k}{2}\left\|z^{(k)} - x^\star\right\|_2^2\right) \tag{A.8}$$

in which case we can combine the bounds to obtain

$$f\left(x^{(k)}\right) - f^\star + \frac{\gamma_k}{2}\left\|z^{(k)} - x^\star\right\|_2^2 \le \left(\prod_{i=0}^{k-1}(1-\theta_i)\right)\left(f\left(x^{(0)}\right) - f^\star + \frac{\gamma_k}{2}\left\|z^{(0)} - x^\star\right\|_2^2\right), \tag{A.9}$$

where we have also used $x^{(0)} = y^{(0)}$ and (A.1) to obtain $x^{(0)} = z^{(0)}$. For completeness, we will show why this is an optimal first-order method. Let $\mu_k = \mu_k^\star = \mu$ and $L_k = L$. If $\gamma_0 \ge \mu$ then using (A.2) we obtain $\gamma_{k+1} \ge \mu$ and $\theta_k \ge \sqrt{\mu/L} = \sqrt{Q^{-1}}$. Simultaneously, we also have $\prod_{i=0}^{k-1}(1-\theta_k) \le \frac{4L}{(2\sqrt{L}+k\sqrt{\gamma_0})^2}$ [30, Lemma 2.2.4], and the bound is then

$$f\left(x^{(k)}\right) - f^\star$$
$$\le \min\left(\left(1-\sqrt{Q^{-1}}\right)^k, \frac{4L}{(2\sqrt{L}+k\sqrt{\gamma_0})^2}\right)\left(f\left(x^{(0)}\right) - f^\star + \frac{\gamma_0}{2}\left\|x^{(0)} - x^\star\right\|_2^2\right). \tag{A.10}$$

This is the optimal convergence rate for the class $\mathcal{F}_{0,L}$ and $\mathcal{F}_{\mu,L}$ simultaneously [28, 30].

Appendix B: Complexity analysis

In this Appendix we prove Theorem 6.1, i.e., we derive the complexity for reaching an ϵ-suboptimal solution for the algorithm *UPN*. The total worst-case complexity is given by (a) the complexity for the worst case number of restarts and (b) the worst-case complexity for a successful termination.

With a slight abuse of notation in this Appendix, $\mu_{k,r}$ denotes the kth iterate in the rth restart stage, and similarly for $L_{k,r}$, $\tilde{L}_{k,r}$, $x^{(k,r)}$, etc. The value $\mu_{0,0}$ is the initial estimate of the strong convexity parameter when no restart has occurred. In the worst case, the heuristic choice in (6.4) never reduces μ_k, such that we have $\mu_{k,r} = \mu_{0,r}$. Then a total of R restarts are required, where

$$\rho_\mu^R\, \mu_{0,0} = \mu_{0,R} \le \mu \quad \Longleftrightarrow \quad R \ge \log(\mu_{0,0}/\mu)/\log(1/\rho_\mu).$$

In the following analysis we shall make use of the relation

$$\exp\left(-\frac{n}{\delta^{-1}-1}\right) \leq (1-\delta)^n \leq \exp\left(-\frac{n}{\delta^{-1}}\right), \quad 0<\delta<1,\ n\geq 0.$$

B.1 Termination complexity

After sufficiently many restarts (at most R), $\mu_{0,r}$ will be sufficiently small in which case (6.9) holds and we obtain

$$\begin{aligned}
&\left\|G_{\tilde{L}_{k+1,r}}\left(x^{(k+1,r)}\right)\right\|_2^2 \\
&\quad\leq \prod_{i=1}^{k}\left(1-\sqrt{\frac{\mu_{i,r}}{L_{i,r}}}\right)\left(\frac{4\tilde{L}_{k+1,r}}{\mu_{k,r}}-\frac{2\tilde{L}_{k+1,r}}{2L_{0,r}}+\frac{2\tilde{L}_{k+1,r}\gamma_{1,r}}{\mu_{k,r}^2}\right)\left\|G_{L_0}\left(x^{(0,r)}\right)\right\|_2^2 \\
&\quad\leq \left(1-\sqrt{\frac{\mu_{k,r}}{L_{k,r}}}\right)^k\left(\frac{4\tilde{L}_{k+1,r}}{\mu_{k,r}}-\frac{2\tilde{L}_{k+1,r}}{2L_{0,r}}+\frac{2\tilde{L}_{k+1,r}\gamma_{1,r}}{\mu_{k,r}^2}\right)\left\|G_{L_{0,r}}\left(x^{(0,r)}\right)\right\|_2^2 \\
&\quad\leq \exp\left(-\frac{k}{\sqrt{L_{k,r}/\mu_{k,r}}}\right)\left(\frac{4\tilde{L}_{k+1,r}}{\mu_{k,r}}-\frac{\tilde{L}_{k+1,r}}{L_{0,r}}+\frac{2\tilde{L}_{k+1,r}\gamma_{1,r}}{\mu_{k,r}^2}\right)\left\|G_{L_{0,r}}\left(x^{(0,r)}\right)\right\|_2^2,
\end{aligned}$$

where we have used $L_{i,r}\leq L_{i+1,r}$ and $\mu_{i,r}\geq\mu_{i+1,r}$. To guarantee $\|G_{\tilde{L}_{k+1,r}}(x^{(k+1,r)})\|_2\leq\bar{\epsilon}$ we require the latter bound to be smaller than $\bar{\epsilon}^2$, i.e.,

$$\begin{aligned}
&\left\|G_{\tilde{L}_{k+1,r}}(x^{(k+1,r)})\right\|_2^2 \\
&\quad\leq \exp\left(-\frac{k}{\sqrt{L_{k,r}/\mu_{k,r}}}\right)\left(\frac{4\tilde{L}_{k+1,r}}{\mu_{k,r}}-\frac{\tilde{L}_{k+1}}{L_{0,r}}+\frac{2\tilde{L}_{k+1,r}\gamma_{1,r}}{\mu_{k,r}^2}\right)\left\|G_{L_{0,r}}(x^{(0,r)})\right\|_2^2\leq\bar{\epsilon}^2.
\end{aligned}$$

Solving for k, we obtain

$$k=\mathcal{O}\left(\sqrt{Q}\log Q\right)+\mathcal{O}\left(\sqrt{Q}\log\bar{\epsilon}^{-1}\right), \tag{B.1}$$

where we have used $\mathcal{O}(\sqrt{L_{k,r}/\mu_{k,r}})=\mathcal{O}(\sqrt{\tilde{L}_{k+1,r}/\mu_{k,r}})=\mathcal{O}(\sqrt{Q})$.

B.2 Restart complexity

How many iterations are needed before we can detect that a restart is needed? The restart detection rule (6.9) gives

$$\begin{aligned}
&\left\|G_{\tilde{L}_{k+1,r}}\left(x^{(k+1,r)}\right)\right\|_2^2 \\
&\quad>\prod_{i=1}^{k}\left(1-\sqrt{\frac{\mu_{i,r}}{L_{i,r}}}\right)\left(\frac{4\tilde{L}_{k+1,r}}{\mu_{k,r}}-\frac{2\tilde{L}_{k+1,r}}{2L_{0,r}}+\frac{2\tilde{L}_{k+1,r}\gamma_{1,r}}{\mu_{k,r}^2}\right)\left\|G_{L_{0,r}}\left(x^{(0,r)}\right)\right\|_2^2
\end{aligned}$$

$$\geq \left(1-\sqrt{\frac{\mu_{1,r}}{L_{1,r}}}\right)^k \left(\frac{4\tilde{L}_{1,r}}{\mu_{1,r}}-\frac{2\tilde{L}_{1,r}}{2L_{0,r}}+\frac{2\tilde{L}_{1,r}\gamma_{1,r}}{\mu_{1,r}^2}\right)\left\|G_{L_{0,r}}\left(x^{(0,r)}\right)\right\|_2^2$$

$$\geq \exp\left(-\frac{k}{\sqrt{L_{1,r}/\mu_{1,r}}-1}\right)\left(\frac{4L_{1,r}}{\mu_{1,r}}-\frac{2L_{1,r}}{2L_{0,r}}+\frac{2L_{1,r}\gamma_{1,r}}{\mu_{1,r}^2}\right)\left\|G_{L_{0,r}}\left(x^{(0,r)}\right)\right\|_2^2,$$

where we have used $L_{i,r} \leq L_{i+1,r}$, $L_{i,r} \leq \tilde{L}_{i+1,r}$ and $\mu_{i,r} \geq \mu_{i+1,r}$. Solving for k, we obtain

$$k > \left(\sqrt{\frac{L_{1,r}}{\mu_{1,r}}}-1\right)\left(\log\left(\frac{4L_{1,r}}{\mu_{1,r}}-\frac{L_{1,r}}{L_{0,r}}+\frac{4\gamma_{1,r}L_{1,r}}{\mu_{1,r}^2}\right)+\log\frac{\|G_{L_{0,r}}(x^{(0,r)})\|_2^2}{\|G_{\tilde{L}_{k+1,r}}(x^{(k+1,r)})\|_2^2}\right). \tag{B.2}$$

Since we do not terminate but restart, we have $\|G_{\tilde{L}_{k+1,r}}(x^{(k+1,r)})\|_2 \geq \bar{\epsilon}$. After r restarts, in order to satisfy (B.2) we must have k of the order

$$\mathcal{O}\left(\sqrt{Q_r}\right)\mathcal{O}(\log Q_r)+\mathcal{O}\left(\sqrt{Q_r}\right)\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right),$$

where

$$Q_r = \mathcal{O}\left(\frac{L_{1,r}}{\mu_{1,r}}\right) = \mathcal{O}\left(\rho_\mu^{R-r}Q\right).$$

The worst-case number of iterations for running R restarts is then given by

$$\begin{aligned}
&\sum_{r=0}^{R}\mathcal{O}\left(\sqrt{Q\rho_\mu^{R-r}}\right)\mathcal{O}\left(\log Q\rho_\mu^{R-r}\right)+\mathcal{O}\left(\sqrt{Q\rho_\mu^{R-r}}\right)\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right)\\
&=\sum_{i=0}^{R}\mathcal{O}\left(\sqrt{Q\rho_\mu^{i}}\right)\mathcal{O}\left(\log Q\rho_\mu^{i}\right)+\mathcal{O}\left(\sqrt{Q\rho_\mu^{i}}\right)\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right)\\
&=\mathcal{O}\left(\sqrt{Q}\right)\left\{\sum_{i=0}^{R}\mathcal{O}\left(\sqrt{\rho_\mu^{i}}\right)\left[\mathcal{O}\left(\log Q\rho_\mu^{i}\right)+\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right)\right]\right\}\\
&=\mathcal{O}\left(\sqrt{Q}\right)\left\{\sum_{i=0}^{R}\mathcal{O}\left(\sqrt{\rho_\mu^{i}}\right)\left[\mathcal{O}(\log Q)+\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right)\right]\right\}\\
&=\mathcal{O}\left(\sqrt{Q}\right)\left\{\mathcal{O}(1)\left[\mathcal{O}(\log Q)+\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right)\right]\right\}\\
&=\mathcal{O}\left(\sqrt{Q}\right)\mathcal{O}(\log Q)+\mathcal{O}\left(\sqrt{Q}\right)\mathcal{O}\left(\log\bar{\epsilon}^{-1}\right)\\
&=\mathcal{O}\left(\sqrt{Q}\log Q\right)+\mathcal{O}\left(\sqrt{Q}\log\bar{\epsilon}^{-1}\right),
\end{aligned} \tag{B.3}$$

where we have used

$$\sum_{i=0}^{R}\mathcal{O}\left(\sqrt{\rho_\mu^{i}}\right)=\sum_{i=0}^{R}\mathcal{O}\left(\sqrt{\rho_\mu}^{\,i}\right)=\mathcal{O}\left(\frac{1-\sqrt{\rho_\mu^{R+1}}}{1-\sqrt{\rho_\mu}}\right)=\mathcal{O}(1).$$

B.3 Total complexity

The total iteration complexity of *UPN* is given by (B.3) plus (B.1):

$$\mathcal{O}\big(\sqrt{Q}\log Q\big) + \mathcal{O}\big(\sqrt{Q}\log\bar{\epsilon}^{-1}\big). \tag{B.4}$$

It is common to write the iteration complexity in terms of reaching an ϵ-suboptimal solution satisfying $f(x) - f^\star \leq \epsilon$. This is different from the stopping criteria $\|G_{\tilde{L}_{k+1,r}}(x^{(k+1,r)})\|_2 \leq \bar{\epsilon}$ or $\|G_{L_{k,r}}(y^{(k,r)})\|_2 \leq \bar{\epsilon}$ used in the *UPN* algorithm. Consequently, we will derive a relation between ϵ and $\bar{\epsilon}$. Using Lemmas 5.1 and 5.2, in case we stop using $\|G_{L_{k,r}}(y^{(k,r)})\|_2 \leq \bar{\epsilon}$ we obtain

$$f\big(x^{(k+1,r)}\big) - f^\star \leq \left(\frac{2}{\mu} - \frac{1}{2L_{k,r}}\right)\big\|G_{L_{k,r}}\big(y^{(k,r)}\big)\big\|_2^2 \leq \frac{2}{\mu}\big\|G_{L_{k,r}}\big(y^{(k,r)}\big)\big\|_2^2 \leq \frac{2}{\mu}\bar{\epsilon}^2,$$

and in case we stop using $\|G_{\tilde{L}_{k+1,r}}(x^{(k+1,r)})\|_2 \leq \bar{\epsilon}$, we obtain

$$\begin{aligned} f\big(\tilde{x}^{(k+1,r)}\big) - f^\star &\leq \left(\frac{2}{\mu} - \frac{1}{2\tilde{L}_{k+1,r}}\right)\big\|G_{\tilde{L}_{k+1,r}}\big(x^{(k+1,r)}\big)\big\|_2^2 \leq \frac{2}{\mu}\big\|G_{\tilde{L}_{k+1,r}}\big(x^{(k+1,r)}\big)\big\|_2^2 \\ &\leq \frac{2}{\mu}\bar{\epsilon}^2. \end{aligned}$$

To return with either $f(\tilde{x}^{(k+1,r)}) - f^\star \leq \epsilon$ or $f(x^{(k+1,r)}) - f^\star \leq \epsilon$ we require the latter bounds to hold and thus select $(2/\mu)\bar{\epsilon}^2 = \epsilon$. The iteration complexity of the algorithm in terms of ϵ is then

$$\begin{aligned} &\mathcal{O}\big(\sqrt{Q}\log Q\big) + \mathcal{O}\big(\sqrt{Q}\log\big((\mu\epsilon)^{-1}\big)\big) \\ &\quad= \mathcal{O}(\sqrt{Q}\log Q) + \mathcal{O}\big(\sqrt{Q}\log\mu^{-1}\big) + \mathcal{O}\big(\sqrt{Q}\log\epsilon^{-1}\big) \\ &\quad= \mathcal{O}(\sqrt{Q}\log Q) + \mathcal{O}\big(\sqrt{Q}\log\epsilon^{-1}\big), \end{aligned}$$

where we have used $\mathcal{O}(1/\mu) = \mathcal{O}(L/\mu) = \mathcal{O}(Q)$.

References

1. Alter, F., Durand, S., Froment, J.: Adapted total variation for artifact free decompression of JPEG images. J. Math. Imaging Vis. **23**, 199–211 (2005)
2. Barzilai, J., Borwein, J.M.: Two-point step size gradient methods. IMA J. Numer. Anal. **8**, 141–148 (1988)
3. Beck, A., Teboulle, M.: Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Trans. Image Process. **18**, 2419–2434 (2009)
4. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**, 183–202 (2009)
5. Becker, S., Bobin, J., Candès, E.J.: NESTA: a fast and accurate first-order method for sparse recovery. SIAM J. Imaging Sci. **4**(1), 1–39 (2011)
6. Becker, S., Candès, E.J., Grant, M.: Templates for convex cone problems with applications to sparse signal recovery. Math. Program. Comput. **3**, 165–218 (2011)

7. Bioucas-Dias, J.M., Figueiredo, M.A.T.: A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. IEEE Trans. Image Process. **16**(12), 2992–3004 (2007)
8. Birgin, E.G., Martínez, J.M., Raydan, M.: Nonmonotone spectral projected gradient methods on convex sets. SIAM J. Optim. **10**, 1196–1211 (2000)
9. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
10. Chambolle, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. **20**, 89–97 (2004)
11. Chambolle, A.: Total variation minimization and a class of binary MRF models. In: Rangarajan, A., Vemuri, B., Yuille, A.L. (eds.) Energy Minimization Methods in Computer Vision and Pattern Recognition. Lecture Notes in Computer Science, vol. 3757, pp. 136–152. Springer, Berlin (2005)
12. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**, 120–145 (2011)
13. Chan, T.F., Golub, G.H., Mulet, P.: A nonlinear primal-dual method for total variation-based image restoration. SIAM J. Sci. Comput. **20**, 1964–1977 (1998)
14. Chan, T.F., Shen, J.: Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods. SIAM, Philadelphia (2005)
15. Combettes, P.L., Luo, J.: An adaptive level set method for nondifferentiable constrained image recovery. IEEE Trans. Image Process. **11**, 1295–1304 (2002)
16. Dahl, J., Hansen, P.C., Jensen, S.H., Jensen, T.L.: Algorithms and software for total variation image reconstruction via first-order methods. Numer. Algorithms **53**, 67–92 (2010)
17. Dai, Y.H., Liao, L.Z.: R-linear convergence of the Barzilai and Borwein gradient method. IMA J. Numer. Anal. **22**, 1–10 (2002)
18. Darbon, J., Sigelle, M.: Image restoration with discrete constrained total variation—Part I: Fast and exact optimization. J. Math. Imaging Vis. **26**, 261–276 (2006)
19. Fletcher, R.: Low storage methods for unconstrained optimization. In: Allgower, E.L., Georg, K. (eds.) Computational Solution of Nonlinear Systems of Equations, pp. 165–179. Am. Math. Soc., Providence (1990)
20. Goldfarb, D., Yin, W.: Second-order cone programming methods for total variation-based image restoration. SIAM J. Sci. Comput. **27**, 622–645 (2005)
21. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton's method. SIAM J. Numer. Anal. **23**, 707–716 (1986)
22. Hansen, P.C.: Discrete Inverse Problems: Insight and Algorithms. SIAM, Philadelphia (2010)
23. Herman, G.T.: Fundamentals of Computerized Tomography: Image Reconstruction from Projections, 2nd edn. Springer, New York (2009)
24. Hintermüller, M., Stadler, G.: An infeasible primal-dual algorithm for total bounded variation-based INF-convolution-type image restoration. SIAM J. Sci. Comput. **28**, 1–23 (2006)
25. Jørgensen, J.H.: Tomobox (2010). www.mathworks.com/matlabcentral/fileexchange/28496-tomobox
26. Kak, A.C., Slaney, M.: Principles of Computerized Tomographic Imaging. SIAM, Philadelphia (2001)
27. Lan, G., Lu, Z., Monteiro, R.D.C.: Primal-dual first-order methods with $O(1/\epsilon)$ iteration-complexity for cone programming. Math. Program., Ser. A **126**(1), 1–29 (2011)
28. Nemirovsky, A.S., Yudin, D.B.: Problem Complexity and Method Efficiency in Optimization. Wiley, New York (1983)
29. Nesterov, Y.: A method for unconstrained convex minimization problem with the rate of convergence $O(1/k^2)$. Sov. Math. Dokl. **269**, 543–547 (1983)
30. Nesterov, Y.: Introductory Lectures on Convex Optimization. Kluwer Academic, Dordrecht (2004)
31. Nesterov, Y.: Smooth minimization of nonsmooth functions. Math. Program., Ser. A **103**, 127–152 (2005)
32. Nesterov, Y.: Gradient methods for minimizing composite objective function (2007). CORE Discussion Paper No 2007076, www.ecore.be/DPs/dp_1191313936.pdf
33. Nolet, G. (ed.): Seismic Tomography with Applications in Global Seismology and Exploration Geophysics. Reidel, Dordrecht (1987)
34. Parrish, R.: getLebedevSphere (2010). www.mathworks.com/matlabcentral/fileexchange/27097-getlebedevsphere
35. Raydan, M.: The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. SIAM J. Optim. **7**, 26–33 (1997)
36. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D **60**, 259–268 (1992)

37. Schabel, M.: 3D Shepp-Logan phantom (2006). www.mathworks.com/matlabcentral/fileexchange/9416-3d-shepp-logan-phantom
38. Tseng, P.: On accelerated proximal gradient methods for convex-concave optimization. Manuscript (2008). www.math.washington.edu/~tseng/papers/apgm.pdf
39. Vandenberghe, L.: Optimization methods for large-scale systems. Lecture Notes (2009). www.ee.ucla.edu/~vandenbe/ee236c.html
40. Vogel, C.R., Oman, M.E.: Iterative methods for total variation denoising. SIAM J. Sci. Comput. **17**, 227–238 (1996)
41. Weiss, P., Blanc-Féraud, L., Aubert, G.: Efficient schemes for total variation minimization under constraints in image processing. SIAM J. Sci. Comput. **31**, 2047–2080 (2009)
42. Zhu, M., Wright, S.J., Chan, T.F.: Duality-based algorithms for total-variation-regularized image restoration. Comput. Optim. Appl. (2008). doi:10.1007/s10589-008-9225-2

Appendix G

Connecting image sparsity and sampling in iterative reconstruction for limited angle X-ray CT

Accepted for the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Lake Tahoe, CA, United States, 2013.

J. S. Jørgensen, E. Y. Sidky and X. Pan

Connecting image sparsity and sampling in iterative reconstruction for limited angle X-ray CT

Jakob S. Jørgensen*, Emil Y. Sidky† and Xiaochuan Pan†
*Department of Applied Mathematics and Computer Science, Technical University of Denmark
Matematiktorvet, bygning 303B, 2800 Kgs. Lyngby, Denmark. Email: jakj@imm.dtu.dk
†Department of Radiology, The University of Chicago
5841 S. Maryland Avenue, Chicago, IL 60637, USA. Email: {sidky,xpan}@uchicago.edu

***Abstract*—A possible quantitative relation between the image sparsity and the number of CT projections views sufficient for accurate reconstruction through 1-norm minimization is investigated empirically. In the setting of full and limited angular range fan-beam and circular cone-beam CT the average number of sufficient views is determined as function of phantom image sparsity over ensembles of randomly generated phantom images. For two phantom classes with different degrees of structure we find a quantitative relation as well as a sharp transition from inaccurate to accurate solution.**

I. Introduction

In the past few years, sparsity-exploiting image reconstruction methods for low-dose computed tomography (CT) have gained interest motivated by the field of compressed sensing (CS) [1], [2]. Numerous studies have demonstrated the potential for accurate reconstruction from a reduced number of measurements both in simulation and on clinical data. As the initial proof-of-concept has been carried out, the excitement over potential large data reduction is developing into new questions on what is missing before these techniques become standard practice [3].

Many factors affect reconstruction quality of sparsity-exploiting methods, including the amount and quality of data, the choice of algorithm and underlying optimization problem and the accuracy with which it is solved as well as the complexity of the test phantom – the topic of the present study. Typically, sparsity-exploiting methods involve many parameters that must be set in just the right way to get a favorable reconstruction, and the large size of realistic CT problems make exhaustive parameter space exploration infeasible. As a result, reconstruction quality of sparsity-exploiting methods remains less understood than for analytical methods.

Recently, we have been studying the role of phantom image complexity for reconstruction quality. Specifically, we have been quantifying the amount of undersampling to expect of a CS-based method in CT [4] and assessing the role of the image sparsity [5], i.e., the number of nonzero pixel values. Image sparsity is a key concept in CS but has to our knowledge not been addressed systematically in CT. In [5], we developed a so-called relative sampling-sparsity (RSS) diagram for investigating a connection between the image sparsity and the number of CT projection views required for accurate reconstruction in the setting of few-view, full angular range CT. The purpose of the present paper is two-fold: to extend the approach to study limited angle problems and to verify the connection between image sparsity and sufficient sampling predicted by small-scale 2-D fan-beam simulations on a 3-D circular cone-beam case.

II. Materials and methods

A. Sparsity-exploiting image reconstruction methods

Sparsity-exploiting methods are motivated by CS-results demonstrating that an image can be reconstructed accurately from a reduced number of measurements [2]. The assumption is that the image is sparse, that is, has a representation with few nonzero coefficients, for example pixel values. For certain discretized forward operators such as partial Fourier matrices and matrices with elements drawn from a Gaussian distribution, theoretical results state how many measurements are needed for *guaranteed* accurate reconstruction of an image of a given sparsity. For system matrices in CT, however, we are unaware of such guarantees, but can investigate a possible connection between sparsity and the number of measurements needed for accurate reconstruction empirically. The establishment of such a connection will provide insight into the amount of undersampling to expect for images of given a given sparsity.

Based on the so-called phase diagram introduced by Donoho and Tanner [6], we proposed in [5] specific for X-ray CT the relative sparsity-sampling (RSS) diagram for studying this connection empirically. Using the diagram we demonstrated the existence of a sharp transition from inaccurate to accurate reconstruction as function of the sparsity and number of measurements for X-ray CT with a 2-D few-view full-angular range scanner configuration. In the present work, we study a limited-angle case using the RSS-diagram.

B. Scanner configuration

We consider a 2-D fan-beam scanner configuration with N_v projections equi-distributed over 360° (full angular-range) or 90° (limited angular-range). The image is restricted to a disk-shaped mask within a $N_\mathrm{side} \times N_\mathrm{side}$ square image, which makes the number of pixels approximately $N = \pi/4 \cdot N_\mathrm{side}^2$. The source-to-center distance is set to $2N_\mathrm{side}$ and the fan-angle to 28.07° for illuminating the disk-shaped image. The detector consists of $N_\mathrm{b} = 2N_\mathrm{side}$ bins, which makes the total number of measurements $M = 2N_\mathrm{side}N_\mathrm{v}$. The line-intersection method

is used for computing X-ray path lengths through the image pixels, each ray yielding an equation of the form

$$b_i = \sum_{j=1}^{N} A_{ij} x_j, \qquad i = 1, \ldots, M, \tag{1}$$

where A_{ij} is the path length of the ith ray through the jth pixel and the system matrix A is of size $M \times N$.

We also consider a 3-D circular cone-beam scanner configuration with circular source trajectory over the same angular ranges. The object is then restricted to a ball-shaped mask within a $N_{\text{side}} \times N_{\text{side}} \times N_{\text{side}}$ cube image and each projection has size $2N_{\text{side}} \times 2N_{\text{side}}$ detector bins.

C. Phantom classes

We use the class of phantoms introduced in [5] called the p-**power** class. The class is originally described in [7] as a background breast tissue model, here followed by thresholding to create zero-valued pixels for obtaining sparse images suited for the experimental design of the present study. The parameter p governs the amount of structure. We can generate random instances of a desired target sparsity from the p-**power** phantom class and refer to a set of such instances as an ensemble. In the present study we consider $p = 0$ and $p = 2$; examples of phantom instances are seen in Fig. 1. The reason for using different phantom classes is to see if sparsity alone can explain the sampling needed for accurate reconstruction or other factors, here structure, play a role as well.

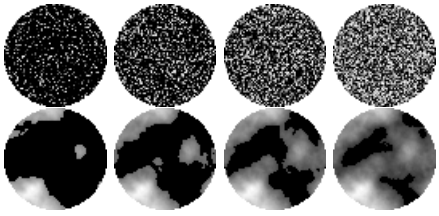


Fig. 1. p-**power** phantom instances. Top, bottom: Structure parameter $p = 0, 2$. Left to right: relative sparsity $\kappa = 0.2, 0.4, 0.6, 0.8$. Gray scale: [0,1].

D. Reconstruction problems and algorithms

For reconstruction, we consider the optimization problem

$$\mathbf{L_1}: \quad x_{\mathbf{L_1}} = \operatorname{argmin} \|x\|_1 \tag{2}$$

$$\text{s.t.} \quad Ax = b. \tag{3}$$

We wish to solve the optimization problem very accurately to avoid false conclusions based on inaccurate solutions. For this purpose, we employ the general-purpose commercial optimization software MOSEK [8], which uses a state-of-the-art primal-dual interior-point method. $\mathbf{L_1}$ can be recast as a linear program (LP), a standard optimization problem to which MOSEK produces a certified primal-dual solution.

For faster solution of the large-scale 3-D problem, we solve instead the problem

$$\mathbf{L_1^\delta}: \quad x_{\mathbf{L_1^\delta}} = \operatorname{argmin} \|x\|_1 \tag{4}$$

$$\text{s.t.} \quad \|Ax - b\|_2^2 \le \delta^2 \tag{5}$$

where the scalar parameter δ acts as a regularization parameter governing the size of the allowed data misfit. For small values of δ and consistent data, the $\mathbf{L_1^\delta}$ solution closely approximates the $\mathbf{L_1}$ solution. We use the Chambolle-Pock algorithm 1 described in [9] with $\delta = 10^{-5}$.

E. Simulation set-up

We create a phantom instance x_{orig} with $N_{\text{side}} = 64$ from one of the p-**power** classes and compute the ideal data $b = Ax_{\text{orig}}$ using different numbers of views, $N_{\text{v}} = 2, 4, 6, \ldots, 32$. We reconstruct by solving $\mathbf{L_1}$ to obtain $x_{\mathbf{L_1}}$. Reconstruction error is measured as the relative 2-norm error to the original, $\|x_{\mathbf{L_1}} - x_{\text{orig}}\|_2 / \|x_{\text{orig}}\|_2$. We accept $x_{\mathbf{L_1}}$ as perfectly recovering x_{orig} if the error is below a threshold of $\epsilon = 10^{-4}$.

With the chosen scanner configuration, we find for both 360° and 90° data that at $N_{\text{v}}^{\text{suf}} = 26$ or more views the system matrix has full column rank, causing x_{orig} to be the unique solution to $Ax = b$. At fewer views, the linear system is underdetermined, with infinitely many solutions and 1-norm minimization is used for selecting a sparse solution. Using $N_{\text{v}}^{\text{suf}}$ as a reference point of having sufficient—or full—sampling, we call $\mu = N_{\text{v}}/N_{\text{v}}^{\text{suf}}$ the relative sampling.

III. Results

A. 2-D fan-beam simulation results: Single phantom instances

First, we wish to demonstrate that $\mathbf{L_1}$ can perfectly recover the original image from 90° data, very similar to what we observed in [5] for 360° data. Fig. 2 shows reconstructions for both 360° and 90° data for $N_{\text{v}} = 6, 8, 10, 12$ of a 0-**power** phantom instance (no structure) and relative sparsity $\kappa = 0.2$. Also shown are difference images with the original to better visualize the transition to recovery. In both cases, we see that at $N_{\text{v}} = 12$ the reconstruction is numerically exact, as the difference images consist only of zeros. Interestingly, $\mathbf{L_1}$ reconstruction of a 0-**power** instance does not appear to be more difficult with the limited angular range of 90°.

We repeat the same experiment with a 2-**power** phantom instance of more structure and show results in Fig. 3. In this case, $N_{\text{v}} = 10$ suffices for accurate reconstruction from the 360° data, while $N_{\text{v}} = 12$ is needed for the 90° data. Apparently, from 360° data the structured phantom is easier to reconstruct than the unstructured, while from 90° data no difference due to structure is seen.

We repeat the experiment for relative sparsity of the 0-**power** phantom instance increased from $\kappa = 0.2$ to 0.4, 0.6 and 0.8. In Fig. 4, reconstruction errors from 360° data are plotted against numbers of views for the four κ-values. The jump to an accurate solution at $N_{\text{v}} = 12$ for $\kappa = 0.2$ from Fig. 2 is recognized. Similar jumps at $N_{\text{v}} = 16, 20, 24$ occur for $\kappa = 0.4, 0.6, 0.8$, and we conclude that the number of

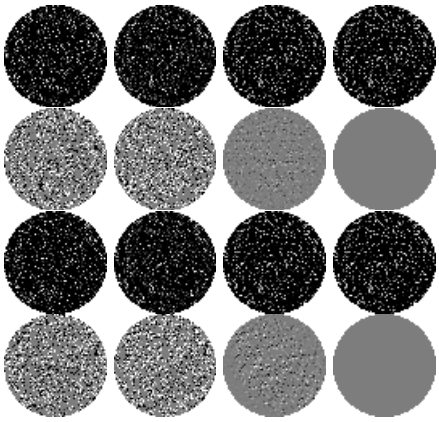


Fig. 2. Left to right: Reconstructions from $N_\mathrm{v} = 6, 8, 10, 12$ views of a 0-**power** phantom instance of relative sparsity $\kappa = 0.2$. 1st/3rd row: 360°/90° data reconstructions. Gray scale: $[0, 1]$. 2nd/4th row: 360°/90° data reconstructions minus original image. Gray scale: $[-0.1, 0.1]$.

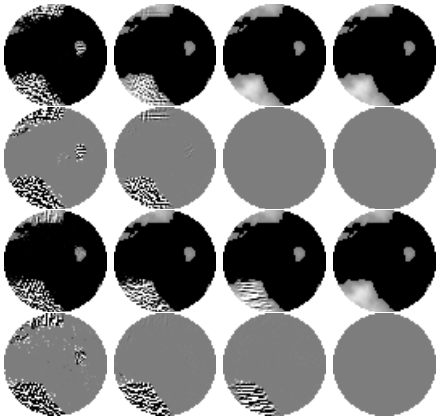


Fig. 3. Same as Fig. 2 for 2-**power** instance of relative sparsity $\kappa = 0.2$.

views needed for accurate $\mathbf{L_1}$-reconstruction appears to grow in a simple way with the relative sparsity κ. Put in another way, we see that images with fewer nonzero pixels admit a larger undersampling relative to the full-sampling reference point of $N_\mathrm{v}^\mathrm{suf} = 26$, as marked by the vertical line in Fig. 4.

B. RSS-diagrams: Multiple phantom instances

A natural question at this point is whether these observations are general or depend on the particular phantom instances used in Figs. 2 and 3. To answer the question, we repeat the experiment for 100 different phantom instances at each

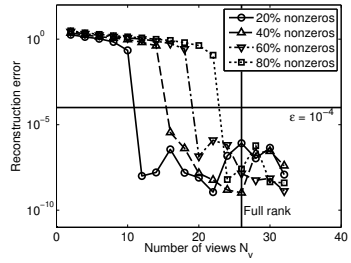


Fig. 4. Reconstruction errors $\|x_{\mathbf{L_1}} - x_\mathrm{orig}\|_2 / \|x_\mathrm{orig}\|_2$ as function of numbers of views N_v for relative sparsity values $\kappa = 0.2, 0.4, 0.6, 0.8$. In all cases, a steep jump from inaccurate to accurate solution is seen and the N_v at which the jump occurs increases with relative sparsity. The vertical line marks the lowest N_v at which the system matrix has full rank.

of the relative sparsity values $\kappa = 0.2, 0.4, 0.6, 0.8$. At each $N_\mathrm{v} = 2, 4, 6, \ldots, 32$ we record the percentage of phantom instances that are reconstructed to within a reconstruction error of $\epsilon = 10^{-4}$. The resulting percentages for the 0-**power** and 2-**power** phantom classes and 360° and 90° data are shown in what we call RSS-diagrams in Fig. 5. Each rectangle represents the percentage of phantoms recovered, ranging from 0% (black) to 100% (white) and shown as function of relative sparsity κ and relative sampling μ. For example, the black bottom left rectangle corresponds to $\kappa = 0.2$ and 2 views, i.e., $\mu = 2/26 \approx 0.08$. In all four cases we recognize the simple connection between relative sparsity and relative sampling sufficient for accurate reconstruction. For the 0-**power** class we observe a very sharp transition from inaccurate to accurate reconstruction in the sense that almost no difference in the relative sampling needed for accurate reconstruction exists among the 100 phantom instances. Furthermore, the RSS-diagrams for 360° and 90° data are identical, which supports our earlier conclusion that $\mathbf{L_1}$-reconstruction of the 0-**power** phantom class is unaffected by the limited angular range.

For the 2-**power** class, the transition from inaccurate to accurate reconstruction is slightly more gradual and for the 360° data occurs about one rectangle (2 views) lower than for the 0-**power** class as well as for the 2-**power** class with 90° data. We conclude that for the more structured phantom class 2-**power**, the limited angular range does make accurate reconstruction with $\mathbf{L_1}$ more difficult.

C. 3-D circular cone-beam simulation results

A practical use of the observed connection between relative sparsity and the relative sampling required for accurate reconstruction is to predict how many views will be needed in other and more difficult-to-simulate scenarios. In [5] we showed that the RSS-diagrams are essentially independent of the image size N_side, so that we can predict sufficient numbers of views at larger pixel arrays based on RSS-diagrams from smaller pixel arrays such as 64×64. Here, we consider predicting the sufficient number of views on a different but related scanner

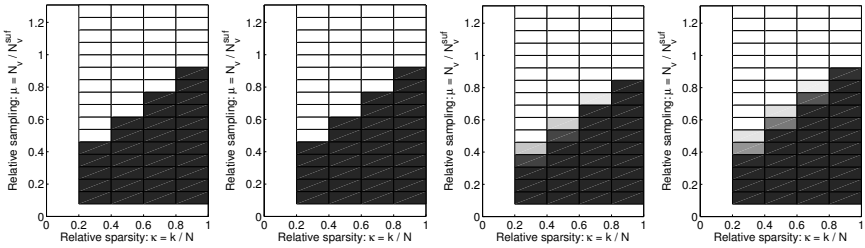


Fig. 5. RSS diagrams: Percentage of accurately reconstructed phantom images as function of relative sparsity and relative sampling. Black=0%, white=100%. Left to right: 0-**power** class with 360° data, 0-**power** class with 90° data, 2-**power** class with 360° data, 2-**power** class with 90° data.

configuration, namely 3-D circular cone-beam. We use a 3-D phantom instance of the 2-**power** class and size $N_{\text{side}} = 32$ with relative sparsity $\kappa = 0.2$. Using the N_{side}-independence of the RSS-diagram we expect at the fifth rectangle from below in the $\kappa = 0.2$ column, which for $N_{\text{side}} = 32$ corresponds to $N_{\text{v}} = 5$, to see a difference between 360° and 90° data. Selected slices of the 3-D $\mathbf{L}_1^{\text{s}}$-reconstructions are shown in Fig. 6 and show excellent agreement with the expectation, as the 360° reconstruction is accurate while the 90° one is not. Interestingly, the central slice, which corresponds precisely to the previous 2-D CT configuration, appears to contradict our expectation as accurate reconstruction is observed in both cases. We explain this by the large degree of sparsity in this plane of the particular phantom instance, because other planes in the 90° reconstruction show prominent errors.

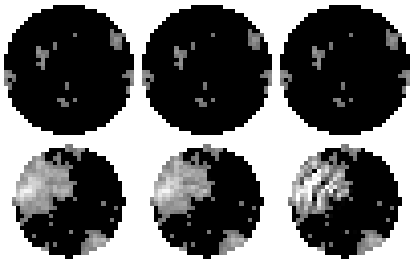


Fig. 6. Top row, left: Central slice (17 of 32 slices, parallel to the plane of the source trajectory) of the 32×32×32 phantom instance from the 2-**power** class of relative sparsity $\kappa = 0.2$. Middle: same slice of 3-D reconstruction from 360° data. Right: same slice of 3-D reconstruction from 90° data. Bottom row: Same for off-central slice (8 of 32). Gray scale: $[0, 1]$.

IV. Discussion and conclusion

The results presented here demonstrate empirically a relation between sparsity of the image to be reconstructed and the average number of fan-beam views required for accurate reconstruction with $\mathbf{L_1}$, both on full angular range and 90° limited angular data. Structured phantoms were found to be accurately reconstructed from slightly fewer views than unstructured phantoms of same sparsity, indicating that image sparsity can only explain some of the variation of the required number of views. The relation found can be used for understanding what undersampling levels to expect when reconstructing sparse images. The RSS-diagram can serve as a tool to investigate such a relation on other sparsity-exploiting methods, e.g., total variation for image gradient sparsity.

Acknowledgment

This work was supported in part by the project CSI: Computational Science in Imaging (The Danish Research Council for Technology and Production Sciences grant 274-07-0065), and in part by The Danish Ministry of Science, Innovation and Higher Education's Elite Research Scholarship. This work was supported in part by NIH R01 grants CA158446, CA120540 and EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

[1] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1289–1306, 2006.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, 2006.

[3] X. Pan, E. Y. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?" *Inverse Prob.*, vol. 25, p. 123009, 2009.

[4] J. S. Jørgensen, E. Y. Sidky, and X. Pan, "Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in X-ray CT." *IEEE Trans. Med. Imag.*, vol. 32, pp. 460–473, 2013.

[5] J. S. Jørgensen, E. Y. Sidky, P. C. Hansen, and X. Pan, "Quantitative study of undersampled recoverability for sparse images in computed tomography," *Submitted*. [Online]. Available: http://arxiv.org/abs/1211.5658

[6] D. Donoho and J. Tanner, "Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing." *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, vol. 367, pp. 4273–4293, 2009.

[7] I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," *Med. Phys.*, vol. 37, pp. 1591–1600, 2010.

[8] MOSEK ApS, "MOSEK Optimization Software (www.mosek.com)," Copenhagen, Denmark. [Online]. Available: www.mosek.com

[9] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Phys. Med. Biol.*, vol. 57, pp. 3065–3091, 2012.

Appendix H

Nonconvex optimization for improved exploitation of gradient sparsity in CT image reconstruction

---

Accepted for the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Lake Tahoe, CA, United States, 2013.

E. Y. Sidky, R. Chartrand, **J. S. Jørgensen** and X. Pan

Nonconvex optimization for improved exploitation of gradient sparsity in CT image reconstruction

Emil Y. Sidky[1], Rick Chartrand[2], Jakob S. Jørgensen[3], and Xiaochuan Pan[1]

***Abstract*—A nonconvex optimization algorithm is developed, which exploits gradient magnitude image (GMI) sparsity for reduction in the projection view angle sampling rate. The algorithm shows greater potential for exploiting GMI sparsity than can be obtained by convex total variation (TV) based optimization. The nonconvex algorithm is demonstrated in simulation with ideal, noiseless data for a 2D fan-beam computed tomography (CT) configuration, and with noisy data for a 3D circular cone-beam CT configuration.**

I. Introduction

Much recent work in iterative image reconstruction in computed tomography (CT) has focused on some form of total variation (TV) minimization, and one of the motivations for employing TV minimization is exploiting sparsity in the gradient magnitude image (GMI) to reduce sampling requirements for the CT system. TV-minimization has been demonstrated, in simulations and with real scanner data, to be effective at allowing for projection view sampling reduction. There is, however, potential to take the sparsity-exploiting principle further, because TV-minimization is an ℓ_1-based convex relaxation of an ideal, nonconvex, sparsity-exploiting optimization based on the ℓ_0-norm. To approach more closely the ℓ_0-based minimization, we develop a GMI sparsity-exploiting algorithm for CT based on an ℓ_p-norm where $p \in (0, 1)$. Section II summarizes the theory and algorithm, and Sec. III shows results based on 2D and 3D CT simulations.

II. Constrained, nonconvex optimization by reweighting

We briefly state the rationale and methods for GMI exploiting CT image reconstruction with nonconvex optimization. We write the CT data model generically as a linear system

$$\mathbf{g} = X\mathbf{f}, \tag{1}$$

where $\mathbf{f}$ is the image vector comprised of voxel coefficients, X is the system matrix generated by some approximation to projection of the voxels, and $\mathbf{g}$ is the data vector containing the estimated projection samples. The model can be applied equally to 2D and 3D geometries, and we note that there are many specific forms to this linear system depending on sampling, image expansion elements, and approximation of continuous fan- or cone-beam projection.

[1]The University of Chicago, Department of Radiology MC-2026, 5841 S. Maryland Avenue, Chicago, IL 60637. Corresponding author: Emil Y. Sidky, E-mail: sidky@uchicago.edu. [2]Theoretical Division, T-5, MS B284, Los Alamos National Laboratory, Los Alamos, NM 87545. [3]Technical University of Denmark, Department of Applied Mathematics and Computer Science, Matematiktorvet, bygning 303B, 2800 Kgs. Lyngby, Denmark.

For the present work, we focus on CT configurations with sparse angular sampling, where the sampling rate is too low for Eq. (1) to have a unique solution. In this situation, there has been much interest in exploiting GMI sparsity of the object to narrow the solution space of Eq. (1) and potentially obtain an accurate reconstruction from under-sampled data. The formulation of this idea results in a nonconvex constrained optimization:

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\| \sqrt{(\partial_x \mathbf{f})^2 + (\partial_y \mathbf{f})^2 + (\partial_z \mathbf{f})^2} \right\|_0 \quad \text{such that } \mathbf{g}_{\text{data}} = X\mathbf{f}, \tag{2}$$

where the argument of the ℓ_0-norm is the voxel-wise magnitude of the image spatial gradient; the linear operators ∂_x, ∂_y, and ∂_z are matrices representing finite differencing in their respective labeled directions; the numerical gradient of the image is formed by, $\nabla\mathbf{f} = [\partial_x\mathbf{f}, \partial_y\mathbf{f}, \partial_z\mathbf{f}]^T$ (2D is obtained by deleting the third component); the ℓ_0-norm counts the number of non-zero components in the argument vector; and $\mathbf{g}_{\text{data}}$ is the available projection data. In words, this optimization seeks the image $\mathbf{f}$ with the lowest GMI sparsity while agreeing exactly with the data.

The optimization problem in Eq. (2) does not lead directly to a practical image reconstruction algorithm, because, as of yet, no large scale solver is available for this problem. Also, the equality constraint, requiring perfect agreement between the available and estimated data, makes no allowance for data inconsistency. In working toward developing a practical image reconstruction algorithm different relaxations of Eq. (2) have been considered. One such relaxation is

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \left\| \sqrt{(\partial_x \mathbf{f})^2 + (\partial_y \mathbf{f})^2 + (\partial_z \mathbf{f})^2} \right\|_p^p \quad \text{such that } \|\mathbf{g}_{\text{data}} - X\mathbf{f}\|_2 \leq \epsilon, \tag{3}$$

where the ℓ_0-norm is replaced by the ℓ_p-norm,

$$\|\mathbf{v}\|_p^p \equiv \sum_i |v_i|^p,$$

and the data equality constraint is relaxed to an inequality constraint with data-error tolerance parameter ϵ. An important strategy, which has been studied extensively in Compressive Sensing [1], is to set $p = 1$, which corresponds to TV-minimization. This, on the one hand, maintains some of the sparsity seeking features of Eq. (2) and, on the other hand, leads to a convex problem, which has convenient features for algorithm development. For example, a local minimizer is a global minimizer in convex optimization.

Another interesting option for GMI sparsity-exploiting image reconstruction is to consider Eq. (3) for $0 < p < 1$. Such a choice for p leads to nonconvex optimization, which can allow for greater sampling reduction than the $p = 1$ case while maintaining highly accurate image reconstruction. These gains intuitively stem from the fact $p < 1$ is closer to the ideal sparsity-exploiting case of $p = 0$; the catch, however, is on the algorithmic side where one has to deal with potential local minima, which are not part of the global solution set. Despite this potential difficulty, practical algorithms based on this nonconvex principle are available [2,3], and gains in sampling reduction for various imaging systems have been reported for both simulated and real data cases. For X-ray tomography, use of this nonconvex strategy has shown promising results [4,5], but the algorithms proposed in those works for CT are only motivated by the optimization problem in Eq. (3) and are not accurate solvers of this problem. An accurate solver is important for theoretical studies of CT image reconstruction with under-sampled data and may also aid in developing algorithms for limited-data tomographic devices.

For CT, one of the barriers to developing an efficient and accurate solver for Eq. (3) in the nonconvex $p < 1$ case, is that it is already challenging to develop such a solver for the convex $p = 1$ case. In order to handle the latter convex, but non-smooth case, we have been interested in an alternate line of optimization problems, where the salient image metrics are written as constraints instead of in an objective function. It is a strategy similar to the set theoretic approach presented in Ref. [6]; the algebraic reconstruction technique (ART) is a specific realization of this strategy; and this type of approach can be useful for nonconvex constraint sets [7]. For the alternate, constraint-based optimization problem there are efficient, large-scale solvers recently available [8,9].

Returning to GMI sparsity-exploiting image reconstruction, we employ an approach developed in Ref. [9] and alter Eq. (3) to the following

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 \text{ such that } \|\mathbf{g}_{\text{data}} - X\mathbf{f}\|_2 \leq \epsilon \text{ and } \left\|\sqrt{(\partial_x \mathbf{f})^2 + (\partial_y \mathbf{f})^2 + (\partial_z \mathbf{f})^2}\right\|_p^p \leq \gamma, \quad (4)$$

which seeks the image $\mathbf{f}$ closest to a prior image $\mathbf{f}_{\text{prior}}$ while respecting constraints on the ℓ_p-norm of the GMI and data-error tolerance. We do not consider, here, the availability of a prior image and set $\mathbf{f}_{\text{prior}} = 0$, keeping this vector only for generality. Consider, first, the case of $p = 1$; the constraint on the GMI becomes a constraint directly on the image TV. Constrained minimization of image TV is known to encourage GMI sparsity. We do not directly minimize TV, rather we independently select parameters γ and ϵ. For sparsity-exploiting image reconstruction, both of these parameters are chosen to have small values: small ϵ forces tight agreement with the data, and small γ encourages GMI sparsity. We note that $\epsilon = 0$ corresponds to a data equality constraint, which may allow no solutions when inconsistencies are present in the data. For $p = 1$, the optimization problem in Eq. (4) is convex and the algorithm presented in Ref. [9] can be used directly to obtain the solution.

For this abstract, we are interested in developing an algorithm for $0 < p < 1$, where the GMI constraint becomes nonconvex. The issue then becomes how to solve Eq. (4) for $p < 1$, because the algorithm in Ref. [9] applies only to convex problems. The approach taken involves approximating Eq. (4) with a convex problem employing a weighted ℓ_1-norm:

$$\mathbf{f}^\circ = \arg\min_{\mathbf{f}} \frac{1}{2}\|\mathbf{f} - \mathbf{f}_{\text{prior}}\|_2^2 \text{ such that } \|\mathbf{g}_{\text{data}} - X\mathbf{f}\|_2 \leq \epsilon \text{ and } \left\|\mathbf{w}\sqrt{(\partial_x \mathbf{f})^2 + (\partial_y \mathbf{f})^2 + (\partial_z \mathbf{f})^2}\right\|_1 \leq \gamma, \quad (5)$$

where the GMI constraint involves only the ℓ_1-norm and a non-negative weight vector $\mathbf{w}$. For a given $\mathbf{w}$ this optimization problem is convex and can be solved efficiently using the algorithm in Ref. [9]. To attack the nonconvex problem, we employ a reweighting technique, where there are two loops: an inner loop where Eq. (5) is solved given parameters γ, ϵ, and weight vector $\mathbf{w}$, and an outer loop where the weight vector is adjusted based on the solution of the inner loop:

$$\mathbf{w} = \left(\sqrt{\eta + (\partial_x \mathbf{f})^2 + (\partial_y \mathbf{f})^2 + (\partial_z \mathbf{f})^2}\right)^{p-1}.$$

The parameter η is needed to prevent the singularity at voxels with zero GMI when $p < 1$. For all simulations in this abstract $\eta = 10^{-6}$. With a reweighting approach, an important question is how accurately does the intermediate weighted problem need to be solved in the inner loop so that overall convergence of the outer loop is attained. It turns out for the present reweighting scheme it suffices to have only one inner iteration. Thus, the complete algorithm is derived from the algorithm in Ref. [9], and the weights are recomputed at every iteration based on the current image estimate $\mathbf{f}$.

III. Results

To demonstrate the new image reconstruction algorithm, we perform two sets of experiments. In the first, we employ the algorithm on ideal, noiseless fan-beam CT data where it is possible to recover the exact image. With this ideal simulation, we demonstrate the potential for angular sampling reduction. In the second simulation, we apply the algorithm to circular, cone-beam CT projections with noise. The purpose of the latter simulation is to demonstrate that the algorithm can indeed by applied to 3D CT, and to illustrate the impact of the nonconvex algorithm on data inconsistency.

A. Ideal fan-beam CT simulation

For the 2D simulation we employ the breast phantom shown in Fig. 1. In the figure, the phantom GMI is also shown, which is seen to have many more zeros than the original phantom. It is this sparsity in the GMI, which we seek to exploit in order to reduce angular sampling. The phantom is discretized on a 128×128 pixel array, which is 20 cm on a side. Only the pixels within the largest inscribed circle are allowed to vary, and pixels outside this 20 cm diameter circle are fixed to zero. The fan-beam CT simulation models an X-ray source 40 cm from the isocenter, and a 80 cm source-to-detector distance.

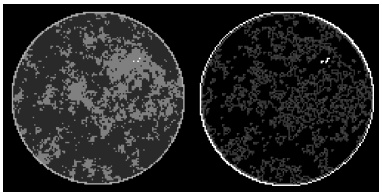
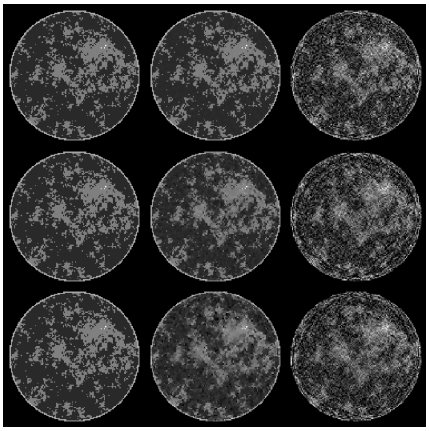


Fig. 1. Left: computerized breast phantom shown in a gray scale window [0.95, 1.25]. Right: gradient magnitude image (GMI), which has greater sparsity than the original phantom.



$p = 0.8$ $p = 1.0$ $p = 2.0$

Fig. 2. Reconstructed images for nonconvex $p = 0.8$, left column, compared with convex $p = 1.0$, middle column, and $p = 2.0$, right column. The number of views covering 360 degrees is 35, 30, and 25 for the top, middle, and bottom rows, respectively. The gray scale window is [0.95, 1.25].

The detector consists of 256 bins in a linear configuration, which is long enough to capture the projection of the 20 cm diameter pixel array. We consider only 360 degree scans, but allow the number of projections to vary.

To illustrate the potential of nonconvex optimization for sparsity-exploiting image reconstruction, we compare solutions of Eq. (4) for $p = 0.8$, $p = 1.0$, and $p = 2.0$. The latter two values lead to a convex problem, which can be solved with the algorithm in Ref. [9], and the first value leads to a nonconvex problem solved by the proposed reweighting algorithm using Eq. (5). For values $p = 1.0$ and $p = 2.0$, we have a direct convergence check, but for the nonconvex case we cannot claim to have found a global solution to Eq. (4). Instead, we can verify that Eq. (5) is solved for the weights $\mathbf{w}$ that are settled upon.

In applying the constraint-based optimization problem in Eq. (4), we need to specify two parameters ϵ and γ. The data used in this simulation are ideal, and accordingly we employ a tight data-error constraint and use a value for ϵ corresponding to an root-mean-square-error (RMSE) of 10^{-5}. For the image TV constraint we set γ to the value of the

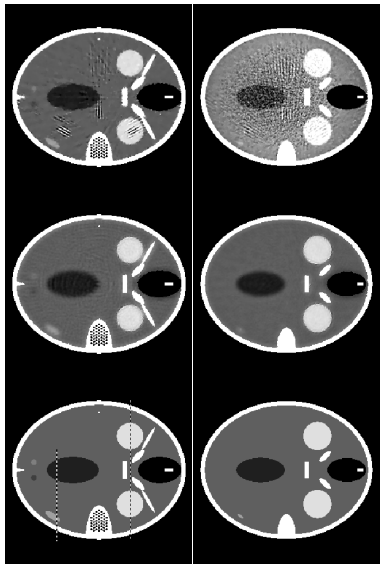


Fig. 3. Images reconstructed from noisy projections of the FORBILD head phantom. The rows show the results for $p = 1.0$, top, $p = 0.8$, middle, and the phantom, bottom in a gray scale window of [1.0425, 1.0625]. The first column shows the midplane, and the second column shows a transaxial plane near the top of the bony structure at the ear. The dashed lines in the phantom midplane slice indicate the locations of the profiles for Figs. 4 and 5.

ℓ_p-norm of the actual phantom GMI to the pth power. We note that in actual application, access to this information is unavailable and selection of γ would need to be based on different image quality metrics. Here, however, we are exploring the theoretical potential of the proposed algorithm.

Shown in Fig. 2, are image reconstruction results for 25, 30, and 35 simulated projections. The $p = 1.0$ case has some potential to reduce angular sampling by exploiting GMI sparsity. This is evident in the comparison with $p = 2.0$, which does not exploit GMI sparsity; the $p = 1.0$ results show visually accurate reconstruction for 35-view projection data while the $p = 2.0$ results do not show accurate reconstruction for any of the projection data sets. The nonconvex $p = 0.8$ results, however, extend the visually accurate reconstruction range down to 25-view projection data.

B. Circular cone-beam CT simulation with noisy projections

For the 3D circular cone-beam CT simulation, we scale up the problem approaching the scale of a realistic volume CT system, and we include noise on the CT projections. The phantom used for this simulation is the FORBILD head phantom, which has many low contrast objects, with gray level variations ranging from 0.25% to 1% of the phantom background, together with complex high-contrast bony structures. This phantom is quite challenging, because even minor streaks from the bony structures can interfere strongly with imaging the low-contrast objects.

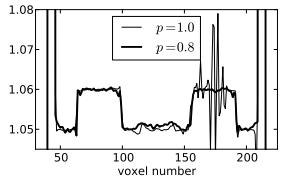


Fig. 4. Profile comparison corresponding to the images in Fig. 3 along a line in the midplane, through the eyes.

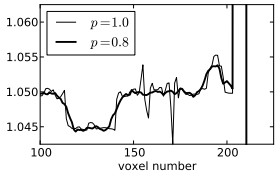


Fig. 5. Profile comparison corresponding to the images in Fig. 3 along a line in the midplane, through the ventricle and subdural hematoma.

The middle section of the head phantom is voxelized in a 256×256×32 volume array, and the projection data simulate 100 projections onto a 512×80 bin flat-panel detector. Noise on the projections is modeled by employing independent 1D Gaussian distributions for each line-integration data value. The mean of each Gaussian distribution is the value of the corresponding line-integration over the phantom, and the standard deviation is taken to be 1% of this mean. The parameters of the simulation are such that it only makes sense to compare algorithms that exploit GMI sparsity, and accordingly we show results from Eq. (4) for $p = 0.8$ and $p = 1.0$. Larger p results in images that are heavily polluted with streak artifacts. For the constraint parameters, we employ an ϵ corresponding to an RMSE of 0.01, and for γ we use the value derived from the test phantom.

For the specified parameters, the image reconstruction results are shown in Fig. 3 together with corresponding slices in the phantom. The gray scale display window is 1% of the phantom complete dynamic range; and streak artifacts are difficult to avoid due to the rapidly oscillating bone structures near the ear at the bottom of the images. The results for $p = 1.0$, in the top row of the figure, show such streaks, even though this value for p does exploit GMI sparsity. The middle row shows results for the nonconvex case of $p = 0.8$, but the streak artifacts are nearly completely removed.

Inspection of the nonconvex results shows a rather interesting behavior in that the image regularization is highly nonuniform. The structures with the contrast of the eyes and greater ($\geq 1\%$ of phantom background) appear to have sharp edges, while the lower contrast structures are visible, yet, are blurred relative to the same structures in the $p = 1.0$ images. This visual impression is borne out quantitatively in vertical profile plots shown through the eyes, in Fig. 4, and through the ventricle and subdural hematoma, in Fig. 5. In the former profile, the nonconvex result has as sharp a transition at the eye border as the convex $p = 1.0$ result without the oscillations from the streaks. The latter lower contrast structures show fewer oscillation for the nonconvex result, but there is also a clear blurring as the transitions at the ventricle and hematoma borders are more gradual for $p = 0.8$ than for $p = 1.0$. This feature of the proposed nonconvex optimization can be understood from inspecting Eq. (5) where we see that the image TV term has a spatially dependent weighting. During the iteration of the nonconvex algorithm the weighting $\mathbf{w}$ evolves in such a way that less weight, and hence less smoothing, is applied to voxels with large gradient-magnitude.

IV. Summary

We have demonstrated GMI sparsity-exploiting image reconstruction by a nonconvex optimization algorithm. Under ideal conditions we have shown that the algorithm is capable of obtaining accurate image recovery with fewer projections than convex TV-based image reconstruction. The algorithm can also be applied to 3D cone-beam CT systems, and preliminary results indicate that the nonconvex algorithm can be effective in controlling steak artifacts resulting from a combination of projection view under-sampling and the presence of complex high-contrast structures.

V. Acknowledgments

This research was supported in part by the U.S. Department of Energy through the LANL/LDRD Program. This work is also part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. This work was also supported in part by NIH R01 grants CA158446, CA120540, and EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

[1] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, 2008.

[2] Rick Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Process. Lett.*, vol. 14, pp. 707–710, 2007.

[3] Rick Chartrand, "Nonconvex splitting for regularized low-rank + sparse decomposition," *IEEE Trans. Signal Process.*, vol. 60, pp. 5810–5819, 2012.

[4] Emil Y. Sidky, Rick Chartrand, and Xiaochuan Pan, "Image reconstruction from few views by non-convex optimization," in *IEEE Nuclear Science Symposium Conference Record*, Honolulu, HI, 2007, pp. 3526 – 3530.

[5] E. Y. Sidky, X. Pan, I. S. Reiser, R. M. Nishikawa, R. H. Moore, and D. B. Kopans, "Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms," *Med. Phys.*, vol. 36, pp. 4920–4932, 2009.

[6] P. L. Combettes, "The foundations of set theoretic estimation," *IEEE Proc.*, vol. 81, pp. 182–208, 1993.

[7] X. Han, J. Bian, E. L. Ritman, E. Y. Sidky, and X. Pan, "Optimization-based reconstruction of sparse images from few-view projections," *Phys. Med. Biol.*, vol. 57, pp. 5245–5274, 2012.

[8] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vision*, vol. 40, pp. 120–145, 2011.

[9] E. Y. Sidky, J. S. Jørgensen, and X. Pan, "First-order convex feasibility algorithms for X-ray CT," *Med. Phys.*, 2013, Accepted. Arxiv preprint (http://arxiv.org/abs/1209.1069).

Appendix I

Sampling conditions for gradient-magnitude sparsity based image reconstruction algorithms

---

In Medical Imaging 2012: Physics of Medical Imaging, editors N. J. Pelc, R. M. Nishikawa and B. R. Whiting, Proc. of SPIE, vol. 8313, p. 831337, San Diego, CA, United States, 2012.
doi:10.1117/12.913307.

E. Y. Sidky, **J. H. Jørgensen** and X. Pan

Sampling conditions for gradient-magnitude sparsity based image reconstruction algorithms

Emil Y Sidky[a], Jakob H. Jørgensen[b], and Xiaochuan Pan[a]

[a]Department of Radiology, The University of Chicago, Chicago, IL
[b]Department of Informatics and Mathematical Modeling, Technical University of Denmark, Kgs. Lyngby, Denmark

ABSTRACT

Image reconstruction from sparse-view data in 2D fan-beam CT is investigated by constrained, total-variation minimization. This optimization problem exploits possible sparsity in the gradient magnitude image (GMI). The investigation is performed in simulation under ideal, noiseless data conditions in order to reveal a possible link between GMI sparsity and the necessary number of projection views for reconstructing an accurate image. Results are shown for two, quite different phantoms of similar GMI sparsity.

Keywords: iterative reconstruction, sparsity, computed tomography

1. INTRODUCTION

Much recent work on iterative image reconstruction in computed tomography (CT) has focused on various forms of constrained, total-variation (TV) minimization.[1–10] These articles were motivated by compressive sensing (CS),[11] where it was suggested that accurate recovery in magnetic resonance imaging from sparse Fourier transform samples may be possible by exploiting sparsity in the gradient-magnitude image (GMI). The idea of exploiting GMI sparsity turns out to be robust, and it can be applied to the CT system.[12] Although much work has shown promising results in applying constrained TV-minimization to image reconstruction in CT both in simulation and with scanner data, the method remains poorly characterized in terms of data requirements and properties of the imaged subject. CS theory does not help with this characterization, because the system matrix employed in CT image reconstruction does not satisfy the conditions of any CS theorems for accurate image recovery.[5]

Accordingly, there are a multitude of fundamental questions having to do with sampling requirements about CT image reconstruction algorithms that exploit GMI-sparsity:
(1) How many views are needed?
(2) What is gained relative to image reconstruction that does not exploit GMI-sparsity?
(3) Should random sampling be employed?
(4) Does constrained, TV-minimization only work on piecewise constant images?
Alternatively, does constrained, TV-minimization lead to stepping artifacts on images that are not piecewise constant?
In this proceedings, we address these questions in a limited way, by performing carefully designed simulations. The simulations are motivated by these questions, but do not provide complete answers.

Section 2 presents the image reconstruction theory for the CT simulations; Sec. 3 shows results of the simulations designed to explore the above questions; and Sec. 4 discusses the simulation results in terms of each of the questions.

2. GMI SPARSITY EXPLOITING IMAGE RECONSTRUCTION FOR CT

The CT data model employed is a linear system:

$$\tilde{g} = X\vec{f}, \tag{1}$$

where $\tilde{g}$ represents the projection data, X is the discrete form of the X-ray transform, and $\vec{f}$ are the image pixel coefficients. For the present work, X is computed by the line-intersection method. In CT it is possible that the

data $\tilde{g}$ are not in the range of X; for the simulations below, however, the data are generated by applying X to test phantoms thereby avoiding the data inconsistency issue. For guaranteeing recovery of the image from this model, the sampling must be such that the condition number of X is finite. To reduce sampling further, prior knowledge on the image function must be exploited. Taking advantage of GMI sparsity is an example of this strategy. To do so involves solving the following equality constrained, TV-minimization problem:

$$\vec{f}^* = \text{argmin } \|\vec{f}\|_{TV} \text{ such that } \tilde{g} = X\vec{f}, \tag{2}$$

where the TV semi-norm is the ℓ_1-norm of the GMI. Basically, this optimization seeks the minimum TV image out of all those that agree perfectly with the data. If the GMI is sparse, this optimization problem can yield perfect recovery, under certain condition, even when X has a non-trivial null-space. Much work in CS aims at establishing the conditions for perfect recovery, but as of yet no results exist that apply to CT.

The effectiveness of exploiting GMI sparsity can be investigated through phantom studies. Performing such studies directly with Eq. (2), however, is difficult algorithmically due to the equality constraint. Instead, we loosen the constraint by introducing a small data-error tolerance ϵ:

$$\ell_2 - TV: \quad \vec{f}^* = \text{argmin } \|\vec{f}\|_{TV} \text{ such that } \|\tilde{g} - X\vec{f}\|_2/\sqrt{N_{\text{data}}} \leq \epsilon, \tag{3}$$

where N_{data} is the total number of measurements; writing the constraint this way allows us to interpret ϵ as a maximum bound on the data root-mean-square-error (RMSE). To solve this problem, we employ an advanced gradient-descent algorithm described in Ref. 13. The parameter ϵ is set to 10^{-5}, a very small value, in the simulations below. Despite this, the increase of the feasible set for $\vec{f}$ can make it difficult to interpret sufficient sampling for accurate recovery with this GMI sparsity-exploiting optimization problem. To aid the interpretation, we employ a quadratic optimization problem which does not exploit GMI sparsity to provide reference images:

$$\ell_2 - \text{magnitude: } \vec{f}^* = \text{argmin } \|\vec{f}\|_2^2 \text{ such that } \|\tilde{g} - X\vec{f}\|_2/\sqrt{N_{\text{data}}} \leq \epsilon. \tag{4}$$

This quadratic minimization is related to Tikhonov regularization and it can be solved with the conjugate gradients algorithm. Sampling conditions for image reconstruction with Eq. (3) and (4) are discussed in greater detail in Ref. 14.

3. 2D SPARSE-VIEW FAN-BEAM CT SIMULATIONS

In Ref. 14 , these sampling conditions are shown for a simulation modeling breast CT. The test phantom in Ref. 14 is a 256×256 image array containing a circular region, where correlated noise is introduced designed to mimic breast fibro-glandular tissue.[15] This phantom has a GMI sparsity of approximately 10,000 non-zero values out of a possible 65,536. For the current work, we employ the same methodology as Ref. 14 on two other phantoms of a similar GMI sparsity level. Although the sparsity is similar, both phantoms are quite different in structure from each other and the breast phantom.

The first "rice" phantom, shown in Fig. 1 consists of many small, thin ellipses, which can also overlap as many as 3 times. Also shown in the figure is the corresponding GMI, which is non-zero at the edges of the ellipses. The other test phantom, shown in Fig. 2, is the FORBILD[16] head phantom with a smooth wavy object included. By itself, the head phantom has few non-zeros in the GMI, but including the wavy object substantially increases the number of GMI non-zeros. The size of this object was set so that the combined phantom also has approximately 10,000 non-zeros in the GMI.

The setting for the CT simulation is 2D fan-beam CT with a 40, and 80, cm source-to-center, and source-to-detector distance, respectively. The test phantoms are discrete 256×256 image arrays, and the data are projected onto a flat detector array of 512 detector bins. The extent of the detector is set so that the inscribed circle of the image array fits exactly onto the detector. The scanning angular range is a full 360°. For the present study, the number of views N_{views} is variable, but the angular intervals between views are constant; namely, this interval is $360°/N_{\text{views}}$. The number of views is varied from 32 to 512 to see at which point the image reconstruction becomes accurate. (Note, that we use the word accurate, here, because the image RMSE will be at best small but non-zero. Were it feasible to compute the solutions for $\epsilon = 0$, we could possibly expect exact

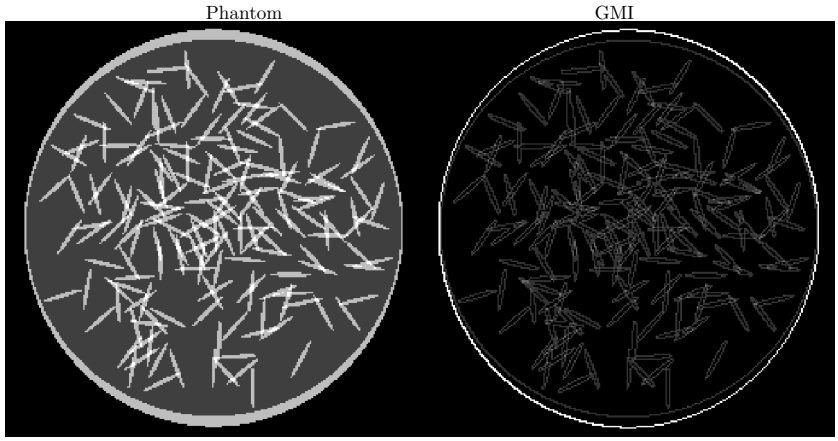


Figure 1. The 256×256 discrete "rice" phantom. Also shown is its gradient magnitude image (GMI). The GMI sparsity is approximately 10,000 out of the 65,536 total pixels.

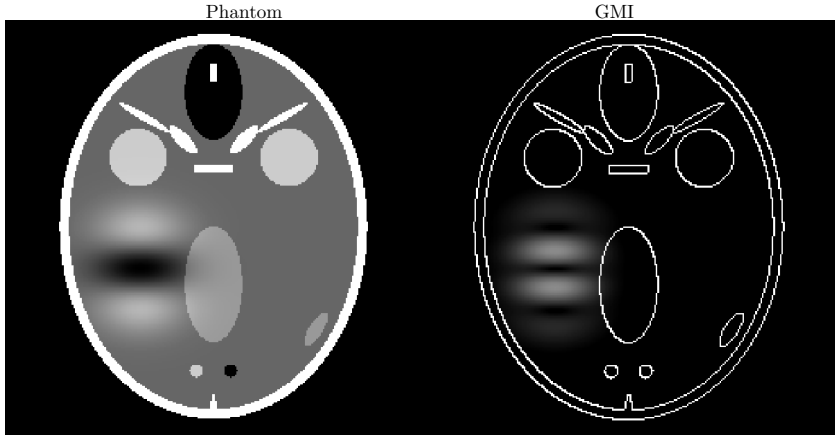


Figure 2. The 256×256 discrete head phantom with a wavy object. The gradient magnitude image (GMI) is shown in a narrow 1% grayscale so that the small values of the wavy object are visible. The GMI sparsity is also approximately 10,000 out of the 65,536 total pixels.

image reconstruction for these idealized simulations.) In order to quantify the image reconstruction accuracy, the image RMSE is computed and plotted as a function of N_{views} with the caveat that image RMSE is a summary metric, which can be insensitive to important image artifacts. Select images are also displayed to show visual image reconstruction accuracy. As stated above, $\epsilon = 10^{-5}$ for both $\ell_2 - TV$ and ℓ_2-magnitude.

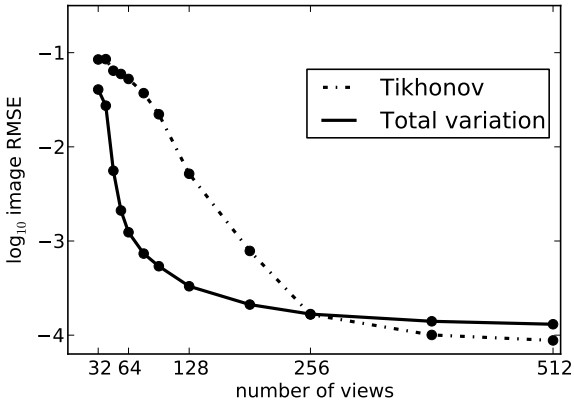


Figure 3. Image RMSEs for images reconstructed of the rice phantom by $\ell_2 - TV$ (referred to as "total variation" in the plot) and ℓ_2-magnitude (referred to as "Tikhonov" in the plot) as a function of N_{views}.

3.1. Results for the rice phantom

The image RMSEs for reconstruction by both $\ell_2 - TV$ and ℓ_2-magnitude using the rice phantom are shown in Fig. 3. Similar results are obtained with the breast phantom and are explained in detail in Ref. 14. We summarize the main points here, briefly. A clear gap in the image RMSE appears between images reconstructed with the GMI sparsity exploiting $\ell_2 - TV$ problem versus the non-sparsity exploiting ℓ_2-magnitude. In fact, the accuracy of reconstruction appears to be high for as low as $N_{\text{views}} = 50$ with $\ell_2 - TV$. This conclusion is corroborated with the selected images shown in Fig. 4. At 48 projection views, X has a large null-space as the total number of samples is actually less than the number of pixel unknowns by a factor of two.

Even when the number of views is large enough that X has no null-space, the image RMSE corresponding to ℓ_2-magnitude is steadily decreasing between N_{views} =101 and 512. This behavior reflects the fact that the condition of X is improving through this range. The gap between the $\ell_2 - TV$ and ℓ_2-magnitude image RMSE curves, in this range, implies that exploiting GMI sparsity can also help stabilize image reconstruction.

One of the ideas of CS is to relate sparsity in the object with the necessary sampling of the sensing system. From this point of view it is interesting to note how many samples are needed for accurate image reconstruction relative to the GMI sparsity. At $N_{\text{views}} = 50$ there are approximately 25,000 samples and the GMI sparsity is 10,000. Thus the sampling to sparsity ratio is approximately 2.5, which is surprisingly good. The theoretical limit for this ratio is 2.0. (To see this, suppose s represents the object sparsity. If $2s-1$ measurements are taken, there will be $2s$-sparse vectors in the null-space of the corresponding sensing matrix X. If there is a $2s$-sparse vector in the null-space of X then two indistinguishable s-sparse vectors can be constructed from the $2s$-sparse vector by separating the coefficients into two equal, disjoint halves.[17]) In light of this limit, a ratio of 2.5 is quite low especially considering that X is far from an ideal matrix for CS.[5]

3.2. Results for the head phantom with wavy object

This phantom is designed to pit against each other two seemingly contradictory ideas about image reconstruction using the TV-norm. Conventional wisdom says that the TV-norm should not be applied to object functions that are not approximately piecewise constant. The CS point-of-view only looks at the identified sparsity in the underlying object function, which is designed to be the same as the rice phantom. From the former point of view, one might expect poor recovery of this phantom or recovery no better than the ℓ_2-magnitude results. From the latter point-of-view, one would expect similar behavior as that of the rice phantom.

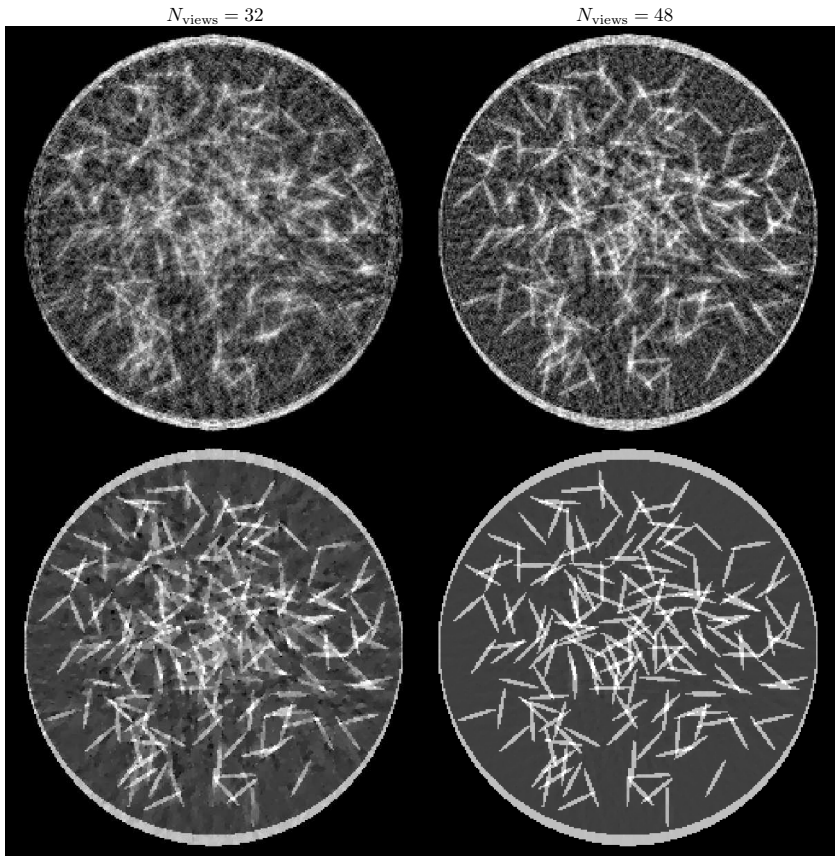


Figure 4. Images of the rice phantom reconstructed by ℓ_2-magnitude (top row) and $\ell_2 - TV$ (bottom row).

The image RMSE results for this phantom are shown in Fig. 5, and surprisingly a different behavior is observed than either of the two anticipated outcomes. The image RMSE is low down to $N_{\text{views}} = 32$ and the selected images of Fig. 6 show accurate image reconstruction with $\ell_2 - TV$ at all N_{views}. One possible explanation for the better than expected results is that the GMI values for the wavy object are much smaller than those corresponding to the edges of the piecewise constant parts of this phantom. With this explanation, accurate image reconstruction is achieved when the number of samples is a factor, possibly 2.5 again, greater than the sparsity of the large GMI values.

The explanation of why the stair-casing artifacts, often seen when TV is applied to smooth non-constant functions, stems from the design of the present simulations. In each case the phantom is a discrete 256×256 grid of pixel values and not a continuous function of the spatial variables. The wavy object approximates a smoothly varying function, but it is in fact piecewise constant as the image function is constant within each pixel.

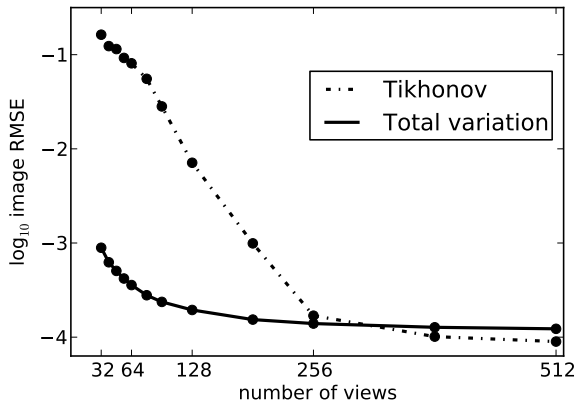


Figure 5. Image RMSEs for images reconstructed of the head phantom by $\ell_2 - TV$ (referred to as "total variation" in the plot) and ℓ_2-magnitude (referred to as "Tikhonov" in the plot) as a function of N_{views}.

4. DISCUSSION OF GMI SPARSITY-EXPLOITING IMAGE RECONSTRUCTION IN CT

With the experience of the limited results above we address the questions about image reconstruction with $\ell_2 - TV$ from the introduction:

(1) How many views are needed?
This depends on the sparsity of the underlying phantom with the rice phantom indicating a possible ratio of 2.5 between the necessary number of samples and the GMI sparsity. The results of the head phantom with a wavy object indicates a more complex rule may be needed that takes into account the relative magnitude of non-zeros in the GMI. Assuming that the ratio of 2.5 holds, denoting the image sparsity by s, and taking the image array as $N \times N$, we choose the number of detector bins to be $2N$ (see Ref. 14). The resulting number of necessary views for accurate image reconstruction by constrained, TV-minimization would then be:

$$N_{\text{views}} \approx 1.25s/N. \tag{5}$$

(2) What is gained relative to image reconstruction that does not exploit GMI-sparsity?
This gain also depends on GMI sparsity. These phantom tests indicate that image reconstruction by $\ell_2 - TV$ may be more stable than ℓ_2-magnitude and it may allow accurate image reconstruction for some system matrices X with a nontrivial null-space.

(3) Should random sampling be employed?
Most CS theorems for exact image reconstruction have been proved for various forms of random sensing matrices.[18] It must be noted, however, those theorems are sufficient conditions which may or may not have a large gap with necessary conditions. The few results shown here, with regular angular-interval sampling, indicate a possible ratio of 2.5 between sampling and GMI sparsity. If this result holds more generally, then there is not a lot of room for improvement. And it is unlikely that randomizing the CT sensing matrix, to the extent allowed by physical constraints, will gain much. We also note that the demonstration illustrated in one of the original CS papers[11] showed sparse FT inversion with a regular sampling pattern.

(4) Does constrained, TV-minimization only work on piecewise constant images?
Alternatively, does constrained, TV-minimization lead to stepping artifacts on images that are not piecewise constant?

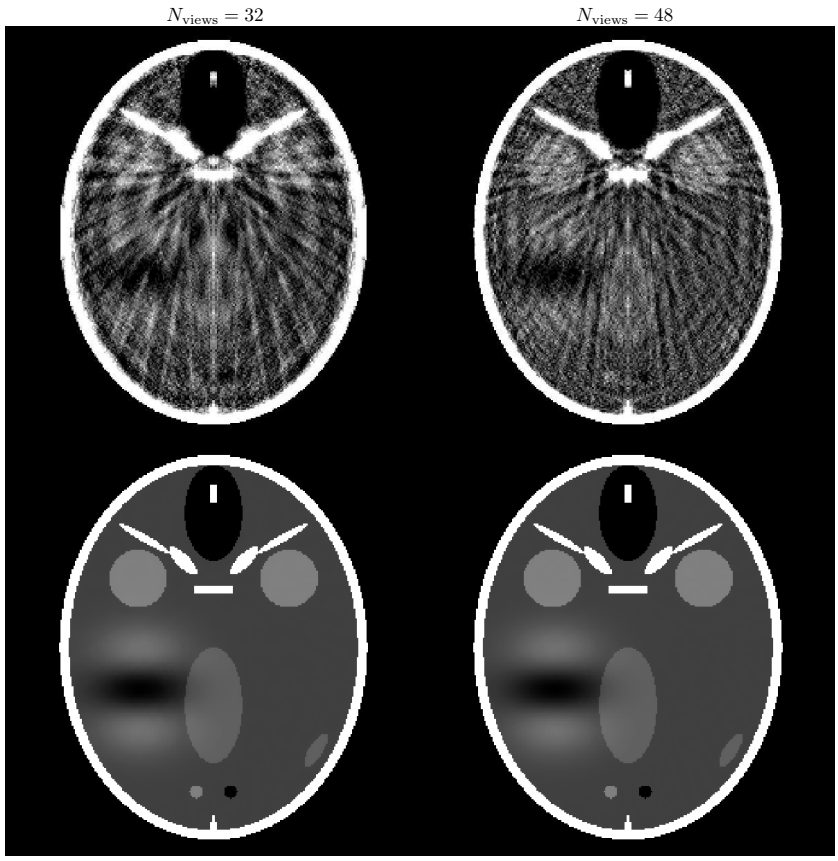


Figure 6. Images of the head phantom with wavy object reconstructed by ℓ_2-magnitude (top row) and $\ell_2 - TV$ (bottom row).

Strictly speaking we did not directly address this question, because this requires a study including data generated from continuous object functions which are smooth and non-constant. In terms of discrete image arrays, it appears that images that closely approximate such functions do not necessarily lead to stepping artifacts. Important factors for accurate image reconstruction are the GMI sparsity, and the number of measurements relative to this sparsity.

The illustrated results are aimed at revealing some of the properties of GMI sparsity exploiting image reconstruction. All the presented results represent best-case scenarios as issues related to various forms of data inconsistency are not considered. Thus obvious extensions of this work would include studies with noisy data or other physical factors such as scatter and beam-hardening. Given that the data model is discrete-to-discrete, having continuous object functions becomes another important factor that will in general lead to data inconsistency. The other important limitation of the work is that it has only been performed on two phantoms and

specifically for the circular fan-beam scanning configuration. It would be of interest to investigate generally applicable relationships between object sparsity and sampling, including more general sampling distributions.

ACKNOWLEDGMENTS

This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. This work was also supported in part by NIH R01 grants CA120540 and EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

REFERENCES

1. J. Song, Q. H. Liu, G. A. Johnson, and C. T. Badea, "Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT," *Med. Phys.* **34**, pp. 4476–4483, 2007.
2. G. H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.* **35**, pp. 660–663, 2008.
3. E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.* **53**, pp. 4777–4807, 2008.
4. X. Pan, E. Y. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?," *Inv. Prob.* **25**, pp. 123009–(1–36), 2009.
5. E. Y. Sidky, M. A. Anastasio, and X. Pan, "Image reconstruction exploiting object sparsity in boundary-enhanced x-ray phase-contrast tomography," *Opt. Express* **18**, pp. 10404–10422, 2010.
6. K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Med. Phys.* **37**, pp. 5113–5125, 2010.
7. X. Jia, Y. Lou, R. Li, W. Y. Song, and S. B. Jiang, "GPU-based fast cone beam CT reconstruction from undersampled and noisy projection data via total variation," *Med. Phys.* **37**, p. 1757, 2010.
8. F. Bergner, T. Berkus, M. Oelhafen, P. Kunz, T. Pan, R. Grimmer, L. Ritschl, and M. Kachelriess, "An investigation of 4D cone-beam CT algorithms for slowly rotating scanners," *Med. Phys.* **37**, pp. 5044–5054, 2010.
9. J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT," *Phys. Med. Biol* **55**, pp. 6575–6599, 2010.
10. X. Han, J. Bian, D. R. Eaker, T. L. Kline, E. Y. Sidky, E. L. Ritman, and X. Pan, "Algorithm-enabled low-dose micro-CT imaging," *IEEE Trans. Med. Imag.* **30**, pp. 606–620, 2011.
11. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**, pp. 489–509, 2006.
12. E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-ray Sci. Tech.* **14**, pp. 119–139, 2006.
13. T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, "Implementation of an optimal first-order method for strongly convex total variation regularization," *BIT, to appear* , 2011.
14. J. H. Jørgensen, E. Y. Sidky, and X. Pan, "Analysis of discrete-to-discrete imaging models for iterative tomographic image reconstruction and compressive sensing," 2011.
arxiv preprint arxiv:1109.0629 (http://arxiv.org/abs/1109.0629).
15. I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," *Med. Phys.* **37**, pp. 1591–1600, 2010.
16. G. Lauritsch and H. Bruder, "FORBILD Head Phantom."
http://www.imp.uni-erlangen.de/phantoms/head/head.html.
17. M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*, Springer, New York, NY, 2010.
18. E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**, pp. 21–30, 2008.

Appendix J

Ensuring convergence in total-variation-based reconstruction for accurate microcalcification imaging in breast X-ray CT

In Proceedings of the 2011 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), pp. 2640–2643, Valencia, Spain, 2011
doi:10.1109/NSSMIC.2011.6152707.
Extended version available from http://arxiv.org/abs/1111.2616.

J. H. Jørgensen, E. Y. Sidky and X. Pan

Ensuring convergence in total-variation-based reconstruction for accurate microcalcification imaging in breast X-ray CT

Jakob H. Jørgensen, *Student Member, IEEE,* Emil Y. Sidky, *Member, IEEE,* and Xiaochuan Pan, *Fellow, IEEE*

***Abstract*—Breast X-ray CT imaging is being considered in screening as an extension to mammography. As a large fraction of the population will be exposed to radiation, low-dose imaging is essential. Iterative image reconstruction based on solving an optimization problem, such as Total-Variation minimization, shows potential for reconstruction from sparse-view data. For iterative methods it is important to ensure convergence to an accurate solution, since important diagnostic image features, such as presence of microcalcifications indicating breast cancer, may not be visible in a non-converged reconstruction, and this can have clinical significance. To prevent excessively long computational times, which is a practical concern for the large image arrays in CT, it is desirable to keep the number of iterations low, while still ensuring a sufficiently accurate reconstruction for the specific imaging task. This motivates the study of accurate convergence criteria for iterative image reconstruction. In simulation studies with a realistic breast phantom with microcalcifications we investigate the issue of ensuring sufficiently converged solution for reliable reconstruction. Our results show that it can be challenging to ensure a sufficiently accurate microcalcification reconstruction, when using standard convergence criteria. In particular, the gray level of the small microcalcifications may not have converged long after the background tissue is reconstructed uniformly. We propose the use of the individual objective function gradient components to better monitor possible regions of non-converged variables. For microcalcifications we find empirically a large correlation between nonzero gradient components and non-converged variables, which occur precisely within the microcalcifications. This supports our claim that gradient components can be used to ensure convergence to a sufficiently accurate reconstruction.**

***Index Terms*—X-ray CT, breast CT, algorithm convergence, total variation, compressed sensing**

I. Introduction

DOSE reduction has gained considerable interest in diagnostic computed tomography (CT) in recent years [1]. The potential to employ CT for screening, where a large population fraction will be exposed to radiation dose and the majority of subjects will be asymptomatic, also motivates the interest in low intensity X-ray CT. Breast CT poses a particularly challenging problem as the total exposure is restricted to the equivalence of two digital mammograms. Such a low X-ray dose can be achieved either by drastically reducing the intensity compared to a diagnostic-quality CT scan, or by reconstruction from sparse-view data.

Total-Variation (TV)-regularized image reconstruction exploits approximate sparsity of the spatial gradient of cross sections of the human body to compensate for reduction in data. TV-reconstructions have been shown to compare favorably with standard Filtered Back Projection from sparse-view data [2], [3]. We are investigating the optimal trade-off between low intensity views and sparse-view data for breast CT by means of TV-reconstruction [4].

The TV-reconstruction is obtained by solving a nonlinear optimization problem. A practical concern is that the extremely large systems in CT, where image arrays of 10^9 voxels are standard, are challenging to solve accurately in acceptable time. Complicating this issue is the fact that clinically relevant features are often very small—occupying only a few voxels. As result both global and pointwise convergence of an iterative reconstruction algorithm may have clinical impact. We demonstrate this issue in the present preliminary investigation, where we examine a realistic simulation of CT for breast cancer screening.

II. Image reconstruction by constrained TV-minimization

We consider TV-regularized image reconstruction in order to exploit gradient sparsity to compensate for the few-view projection data. The present study works with the discrete-to-discrete imaging model, $Au = b$, see [5]. For reconstruction we consider the minimization problem

$$u_{\mathrm{TV}} = \underset{u}{\operatorname{argmin}}\, f(u), \tag{1}$$

where

$$f(u) = \|Au - b\|_1 + \lambda \|u\|_{\mathrm{TV}} \tag{2}$$

and

$$\|u\|_{\mathrm{TV}} = \sum_j \|D_j u\|_2 \tag{3}$$

and D_j is a forward difference approximation to the image gradient at pixel j.

Manuscript received November 8, 2011. This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. This work was supported in part by NIH R01 grants CA120540, EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

J. H. Jørgensen is with the Department of Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Bygning 321, 2800 Kongens Lyngby, Denmark (e-mail: jakj@imm.dtu.dk).

E. Y. Sidky and X. Pan are with the Department of Radiology, University of Chicago, 5841 S. Maryland Ave., Chicago IL, 60637, USA (e-mail: {sidky,xpan}@uchicago.edu).

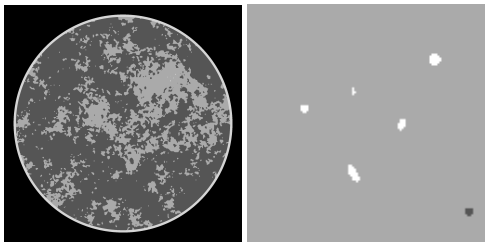


Fig. 1. Left: Original full breast phantom, 2048^2 pixels. Right: 120^2 pixel region of interest around simulated microcalcifications. Gray level window: $[0.9, 1.2]$. The microcalcifications are located within the largest region of fibro-glandular tissue in the upper right quarter.

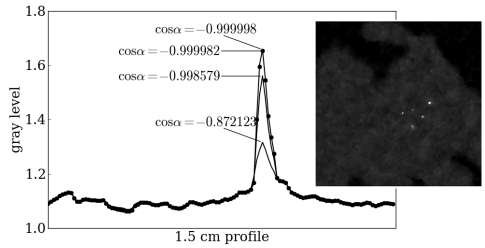


Fig. 2. Profiles through single microcalcification for reconstructions increasingly close to satisfying optimality condition $\cos\alpha = -1$. Inset: 400^2 pixel region of interest of full 2048^2 pixel reconstruction for $\cos\alpha = -0.999998$.

Instead of the more commonly used ℓ_2 norm for measuring data fidelity we use the ℓ_1 norm. TV-regularized ℓ_2 norm minimization is known to be contrast-reducing, in particular for objects of small scale [6], such as microcalcifications. ℓ_1 minimization does not remove this problem, but tends to reduce it [7].

Both terms in (2) are non-differentiable, and in order to apply standard gradient-based optimization algorithms we apply the standard smoothing trick of the replacements:

$$\sum_j \sqrt{\|D_j u\|_2^2 + \epsilon} \qquad \text{replaces} \qquad \|u\|_{\mathrm{TV}}. \tag{4}$$

$$\sum_i \sqrt{|(Au)_i - b_i|^2 + \epsilon} \qquad \text{replaces} \qquad \|Au - b\|_1. \tag{5}$$

In our simulations we use $\epsilon = 10^{-4}$, which we found sufficiently small to prevent any change in visual appearance of the reconstructed image compared to using $\epsilon = 0$.

An important question is how well a TV reconstrution is capable of reproducing the salient image features, such as microcalcifications in the present case. Numerous studies demonstrate that of TV-reconstruction can produce clinically useful reconstructions, see e.g. [2], [3].

Our main question of interest in the present work arises when using an iterative algorithm to solve the TV minimization problem: When can we reliably stop iterating and accept the computed solution as a good approximation of the true minimizer to (2)? We consider here the $\cos\alpha$ optimality condition [3], which says that at the minimizer we have $\cos\alpha = -1$, where α is the angle between the gradients of each of the two terms in (2).

For solving (1) we use a convergent, gradient-based optimization algorithm, which is optimal in a certain sense, see [8]. The algorithm was developed for minimization of the TV-regularized ℓ_2 data fidelity, but is applicable to any smooth objective function, and we have found that it works well for solving (the smoothed version of) the problem in (1).

III. Breast CT model

Breast CT imaging is being considered as a potential addition to mammography in screening for early-stage diagnosing of breast cancer. One particular indicator of breast cancer is formation of *microcalcifications*—very small, highly attenuating calcium deposits. For screening, low-dose imaging is pertinent to minimize accumulated X-ray dose, while accurate and reliable microcalcification shape and attenuation reconstruction is crucial for precise diagnosing.

In the present work we use the breast phantom from [9] discretized on a 2048^2 pixel grid. We include a simulated cluster of microcalcifications, also discretized. Gray values in units of water attenuation are: fat 1.00; fibro-glandular tissue: 1.10; skin: 1.15 and microcalcifications: $1.80 - 2.10$. The phantom is shown in Fig. 1 along with a 120^2 pixel region of interest (ROI) around the simulated cluster of microcalcifications.

IV. Numerical simulation

We wish to demonstrate that the TV-reconstruction is subject to *non-uniform convergence*; more specifically that the pixel values in the microcalcifications converge much more slowly than the background pixel values.

Our concern about non-uniform convergence arises from two facts: First, detecting non-uniform convergence can be very challenging as we will demonstrate. Second, if we are not aware of non-uniform convergence, we risk accepting a solution which is not yet converged everywhere. Such a reconstruction has much lower contrast than the true TV-solution, which will make it difficult to spot the microcalcifications. This can lead us to the, incorrect, conclusion that the TV-solution is not capable of reproducing microcalcifications faithfully, but in fact the lack of contrast in the reconstruction was a result of accepting a too early iterate returned by the iterative solver and not because of the TV-minimization problem itself.

We generate noise-free 64-view, 1024-detector-bin fan-beam data by forward projection (using a line intersection-based ray-driven projector) of the original discrete 2048^2 pixelized phantom with microcalcifications. In Fig. 2 we show four profiles through a microcalcification from reconstructions at $\cos\alpha = -0.872123$, -0.998579, -0.999982, and -0.999998, i.e., increasingly close to satisfying the optimality condition $\cos\alpha = -1$. We also show a region of interest around the final reconstruction, demonstrating that the microcalcifications can be reconstructed by TV-reconstruction. From previous investigations, although without objects of similarly small

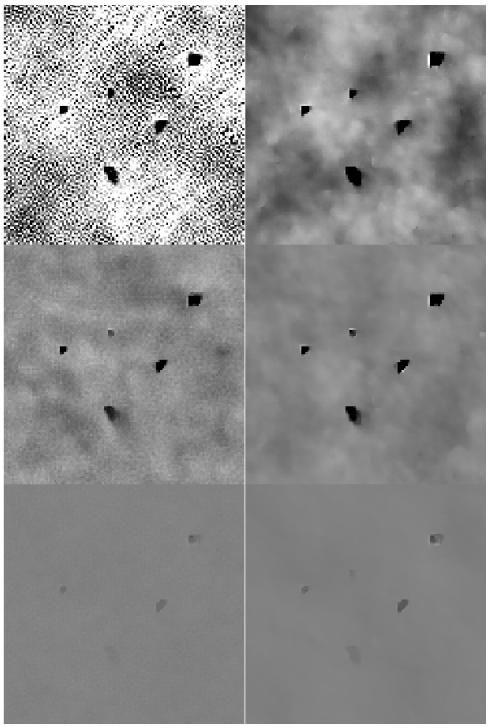


Fig. 3. Left column: 120^2 pixel ROI gradient components, grayscale window: $[-0.0001, 0.0001]$. Right column: 120^2 pixel ROI absolute difference images to most accurate reconstruction (for $\cos\alpha = -0.999998$), grayscale window: $[-0.02, 0.02]$. Top to bottom: $\cos\alpha = -0.872123$, -0.998579, -0.999982.

scale, we have the experience that a $\cos\alpha$ of -0.8 or even -0.5 is sufficient [3] for achieving useful reconstructions from real scanner data.

For the present microcalcification simulation we observe a non-uniform convergence across the image, in the sense that while no change in the background is seen after $\cos\alpha = -0.872123$, it takes until at least $\cos\alpha = -0.999982$ before the microcalcification-pixels reach convergence. We conclude that if we simply use our empirical target value of $\cos\alpha = -0.8$ we will fail in reconstructing microcalcifications sufficiently well. Furthermore, it is likely that a sufficient value of $\cos\alpha$ be dependent on the size and contrast of the microcalcifications as well as other parameters such as the the discretization and λ, which makes it difficult to decide an appropriate target value.

V. Gradient components

As a first step towards a more reliable convergence criterion we wish to point out a connection that can possibly exploited. The considered $\cos\alpha$ convergence criteria involves the gradient of the objective function f, and so do other standard optimality conditions [10]. However, as we saw, it is not clear *how close* to -1 we must $\cos\alpha$ to be, for ensuring that all pixel values have reached convergence. We conclude that the $\cos\alpha$-criterion is not sensitive enough to detect the few pixel values that have still not settled. We believe this is due to computing a single number ($\cos\alpha$) from the full gradient for comparing with the optimal value of -1, thereby "averaging out" the differences between the individual components of the gradient. Many small gradient components will tend to hide the presence of a few larger ones.

We propose instead to replace the use of a single number convergence criterion such as the $\cos\alpha$ by monitoring the objective function gradient displayed as an image: For the jth pixel of the image x and the objective function to be minimized $f(u)$ we refer to $(\nabla f(u))_j$ as the jth gradient component. A necessary condition, and part of the KKT optimality conditions [10], is that all gradient components are zero. We emphasize that it is perfectly possible for a gradient component to be zero even though the corresponding variable has not converged, even for a convex objective function. An example is a convex quadratic of only two variables, having different length semi axes which are not aligned with the coordinate axes. The minimizer is at the origin but there are two straight lines along which either of the gradient components are zero, but not both. Empirically, however, we have observed strong evidence of good correlation between non-zero gradient components and the remaining non-converged pixel values in the image.

For the microcalcification simulation we show ROIs of the gradient components as images in Fig. 3. Since we do not know the true solution, we use the $\cos\alpha = -0.999998$ reconstruction as an approximation to the true solution, and we show difference images of the reconstructions to the reference image. The reason for comparing to the true solution is that we want to determine whether the gradient components, which are readily available from a given iterate, can be used to predict regions of unconverged variables, which are, of course, unknown at any given point during the iterations.

We observe a highly non-uniform nonzero gradient component pattern for the least accurate reconstruction, with large (negative) components exactly at the microcalcifications. The gradient components are negative, which agress with the variables still growing as seen in the profile in Fig. 2. For the more accurate reconstructions the microcalcification gradient components remain distinct while their intensities vanish. We observe a large correlation with the non-converged variables of the difference images, indicating a close connection. This suggests the possibility for ensuring local convergence in the microcalcifications by means of monitoring the gradient components.

We find that the gradient components on the the microcalcifications in Fig. 3 bottom are sufficiently small that we are confident that the image has converged, opposed to basing the convergence on a $\cos\alpha$ of -0.999982, which we have no straightforward method for judging whether is "close enough" to -1.

VI. Discussion

We are investigating strategies other than visual inspection of the gradient components for a quantitative convergence criterion. For instance by forcing $\max_j |(\nabla f(u))_j|$ below an appropriately chosen threshold ϵ, all gradient components will be smaller than ϵ, thereby ensuring global convergence. When applying a single number based convergence criterion such the $\cos\alpha$-criterion, the fact that the majority of the variables are at optimum can conceal by averaging out the contributions from the few variables that are not. The rationale in forcing all gradient components below ϵ is that small areas of non-convergent varibles will prevent termination of the algorithm. A different approach would be to exploit the spatial structure in the nonzero gradient components, e.g. by not terminating iterations until no spatial correlation is present.

The use of the objective gradient in a convergence criterion is well-known, at least the use of the norm of the gradient. Explicit use of the individual gradient components for monitoring local convergence for small objects such as microcalcifications has—to the best of our knowledge—not been studied before.

VII. Conclusion and future work

We have conducted a preliminary investigation of non-uniform convergence for reconstruction of microcalcifications in breast CT. We saw that it is potentially difficult to ensure a sufficiently converged solution simply by use of a convergence criterion such as the $\cos\alpha$-criterion, due to non-uniform convergence caused by the small size of the microcalcifications.

Accepting a reconstruction which is not globally converged may have clinical significance, for instance, as in the example given, by providing insufficient contrast for detecting the microcalcifications. We demonstrated that the nonzero gradient components can be used to monitor the regions of non-converged variables and thereby prevent termination of the optimization algorithm before global convergence is reached.

Interesting directions for future work include to develop a quantitative convergence criterion based on gradient components, as well as to investigate the use in other optimization based reconstruction techniques besides TV-reconstruction.

References

[1] C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, "Strategies for reducing radiation dose in ct," *RADIOLOGIC CLINICS OF NORTH AMERICA*, vol. 47, pp. 27–40, 2009.

[2] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *JOURNAL OF X-RAY SCIENCE AND TECHNOLOGY*, vol. 14, no. 2, pp. 119–139, 2006.

[3] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *PHYSICS IN MEDICINE AND BIOLOGY*, vol. 53, no. 17, pp. 4777–4807, 2008.

[4] J. H. Jørgensen, P. C. Hansen, E. Y. Sidky, I. S. Reiser, and X. Pan, "Toward optimal X-ray flux utilization in breast CT," in *Proceedings of the 11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Potsdam, Germany, 2011, available from http://arxiv.org/abs/1104.1588.

[5] H. H. Barrett and K. J. Myers, *Foundations of Image Science*. Hoboken, NJ: John Wiley & Sons, 2004.

[6] D. Strong and T. Chan, "Edge-preserving and scale-dependent properties of total variation regularization," *INVERSE PROBLEMS*, vol. 19, no. 6, pp. S165–S187, 2003.

[7] T. F. Chan and S. Esedoglu, "Aspects of total variation regularized L-1 function approximation," *SIAM JOURNAL ON APPLIED MATHEMATICS*, vol. 65, no. 5, pp. 1817–1837, 2005.

[8] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, "Implementation of an optimal first-order method for strongly convex total variation regularization," *BIT, to appear*, 2011, preprint available from: http://arxiv.org/abs/1105.3723.

[9] I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," *MEDICAL PHYSICS*, vol. 37, pp. 1591–1600, 2010.

[10] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York: Springer, 2006.

Appendix K

Accelerated gradient methods for total-variation-based CT image reconstruction

In Proceedings of the 11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 435–438, Potsdam, Germany, 2011.
Proceedings available from: http://www.fully3d.org.

J. H. Jørgensen, T. L. Jensen, P. C. Hansen, S. H. Jensen, E. Y. Sidky and X. Pan

Accelerated gradient methods for total-variation-based CT image reconstruction

Jakob H. Jørgensen, Tobias L. Jensen, Per Christian Hansen, Søren H. Jensen, Emil Y. Sidky, and Xiaochuan Pan

***Abstract*—Total-variation (TV)-based CT image reconstruction has shown experimentally to be capable of producing accurate reconstructions from sparse-view data. In particular TV-based reconstruction is well suited for images with piecewise nearly constant regions. Computationally, however, TV-based reconstruction is demanding, especially for 3D imaging, and the reconstruction from clinical data sets is far from being close to real-time. This is undesirable from a clinical perspective, and thus there is an incentive to accelerate the solution of the underlying optimization problem.**

The TV reconstruction can in principle be found by any optimization method, but in practice the large scale of the systems arising in CT image reconstruction preclude the use of memory-intensive methods such as Newton's method. The simple gradient method has much lower memory requirements, but exhibits prohibitively slow convergence.

In the present work we address the question of how to reduce the number of gradient method iterations needed to achieve a high-accuracy TV reconstruction. We consider the use of two accelerated gradient-based methods, GPBB and UPN, to solve the 3D-TV minimization problem in CT image reconstruction. The former incorporates several heuristics from the optimization literature such as Barzilai-Borwein (BB) step size selection and nonmonotone line search. The latter uses a cleverly chosen sequence of auxiliary points to achieve a better convergence rate.

The methods are memory efficient and equipped with a stopping criterion to ensure that the TV reconstruction has indeed been found. An implementation of the methods (in C with interface to Matlab) is available for download from `http://www2.imm.dtu.dk/~pch/TVReg/`**.**

We compare the proposed methods with the standard gradient method, applied to a 3D test problem with synthetic few-view data. We find experimentally that for realistic parameters the proposed methods significantly outperform the standard gradient method.

***Index Terms*—Total-variation, Gradient-based optimization, Strong convexity, Algorithms**

I. Introduction

Algorithm development for image reconstruction from incomplete data has experienced renewed interest in the past years. Incomplete data arises for instance in case of a small number of views, and the development of algorithms for incomplete data thus has the potential for a reduction in imaging time and the delivered dosage.

Jakob H. Jørgensen and Per Christian Hansen are with the Department of Informatics and Mathematical Modelling, Technical University of Denmark, Richard Petersens Plads, Building 321, 2800 Kgs. Lyngby, Denmark. Corresponding author: Jakob H. Jørgensen, E-mail: jakj@imm.dtu.dk.

Tobias L. Jensen and Søren H. Jensen are with the Department of Electronic Systems, Aalborg University, Niels Jernesvej 12, 9220 Aalborg Ø, Denmark.

Emil Y. Sidky and Xiaochuan Pan are with the Department of Radiology, The University of Chicago, 5841 S. Maryland Avenue, Chicago, IL 60637, USA.

Total-varation (TV)-based image reconstruction is a promising direction, as experiments have documented the potential for accurate image reconstruction under conditions such as few-view and limited-angle data, see e.g. [1].

However, it also known that it is difficult to design fast algorithms for obtaining exact TV reconstructions due to nonlinearity and nonsmoothness of the underlying optimization problem. Many different approaches have been developed, such as time marching [2], fixed-point iteration [3], and various minimization-based methods such as sub-gradient methods [4], second-order cone programming (SOCP) [5], and duality-based methods [6], [7] – but for large-scale applications such as CT image reconstruction the computational burden is still unacceptable. As a consequence heuristic and much faster techniques such as the one in [1] for approximating the TV solution have been developed. In such inaccurate, but efficient, TV-minimization solvers the resulting image depends on several algorithm parameters, which introduces an unavoidable variability. In contrast, for the accurate TV algorithms considered here, the resulting image can be considered dependent only on the parameters of the optimization problem.

In this work we present two accelerated gradient-based optimization methods that are capable of computing the TV reconstruction of 3D volumes to within an accuracy specified by the user.

II. Theory

A. Total-variation-based image reconstruction

In this work we consider total-variation (TV)-based image reconstruction for computed tomography. The 3D reconstruction is represented by the vector x^* which is the solution to the minimization problem

$$x^* = \operatorname*{argmin}_{x \in \mathcal{Q}} \ \phi(x), \quad \phi(x) = \frac{1}{2}\|A\,x - b\|_2^2 + \alpha\|x\|_{\mathrm{TV}}. \quad (1)$$

Here, x is the unknown image, $\mathcal{Q}$ is the set of feasible x, A is the system matrix, b is the projection data stacked into a column vector, and α is the TV regularization parameter specifying the relative weighting between the fidelity term and the TV term. $\|x\|_{\mathrm{TV}}$ is the discrete total-variation of x,

$$\|x\|_{\mathrm{TV}} = \sum_{j=1}^{N} \|D_j x\|_2, \quad (2)$$

where N is the number of voxels and D_j is the forward difference approximation to the gradient at voxel j.

B. Smooth and strongly convex functions

We recall that a continuously differentiable function f is convex if

$$f(x) \geq f(y) + \nabla f(y)^T (x - y) \tag{3}$$

for all $x,\ y \in \mathcal{Q}$. A stronger notion of convexity is *strong convexity*: f is said to be *strongly convex* with *strong convexity parameter* μ if there exists a $\mu \geq 0$ such that

$$f(x) \geq f(y) + \nabla f(y)^T (x - y) + \frac{1}{2}\mu \|x - y\|_2^2 \tag{4}$$

for all $x,\ y \in \mathcal{Q}$. Furthermore, f has *Lipschitz continuous* gradient with *Lipschitz constant* L, if

$$f(x) \leq f(y) + \nabla f(y)^T (x - y) + \frac{1}{2} L \|x - y\|_2^2 \tag{5}$$

for all $x,\ y \in \mathcal{Q}$. The ratio μ/L is important for the convergence rate of gradient methods we will consider.

The problem (1) can be shown [8] to be strongly convex and have Lipschitz continuous gradient in the case where A specifies a full-rank overdetermined linear system. In the rank deficient or underdetermined case, which occurs for instance for few-view data, the strong convexity assumption is violated. However, as we shall see, this turns out not to pose a problem for the gradient methods we consider.

III. Algorithms

A. Gradient projection methods

The optimization problem (1) can, in principle, be solved by use of a simple gradient projection (**GP**) method

$$x^{(k+1)} = P_{\mathcal{Q}} \left(x^{(k)} - \theta_k \nabla f(x^{(k)}) \right), \quad k = 0, 1, 2, \ldots, \tag{6}$$

where $P_{\mathcal{Q}}$ denotes projection onto the set $\mathcal{Q}$ of feasible x, and θ_k is the step size at the kth step. The worst-case convergence rate of **GP** with $\mu > 0$ and constant step size is

$$f(x^{(k)}) - f^\star \leq \left(1 - \frac{\mu}{L}\right)^k \cdot C_{\mathrm{GP}}, \tag{7}$$

where C_{GP} is a constant [9, §7.1.4].

For large-scale imaging modalities, such as CT, this slow convergence renders the simple gradient method impractical. On the other hand the simplicity and the low memory requirements of the gradient method remain attractive. Various modifications have been suggested in the optimization literature. For instance, a significant acceleration is often observed empirically if the gradient method is equipped with a Barzilai-Borwein (BB) step size strategy and a nonmonotone line search [10], [11], [12], [13], [14], see Algorithm 1: **GPBB** for a pseudo-code. Empirically we have found $K = 2$ and $\sigma = 0.1$ to be satisfactory parameter choices. However, it remains unproven that **GPBB** achieves a better worst-case convergence rate than (7).

Algorithm 1: **GPBB**

input : $x^{(0)}$, K
output: $x^{(k+1)}$

$\theta_0 = 1$;
for $k = 0, 1, 2, \ldots$ **do**
  // BB strategy
  if $k > 0$ **then**
    $\theta_k \leftarrow \frac{\|x^{(k)} - x^{(k-1)}\|_2^2}{\langle x^{(k)} - x^{(k-1)}, \nabla f(x^{(k)}) - \nabla f(x^{(k-1)}) \rangle}$;
  $\beta \leftarrow 0.95$;
  $\bar{x} \leftarrow P_{\mathcal{Q}}(x^{(k)} - \beta\theta_k \nabla f(x^{(k)}))$;
  $\hat{f} \leftarrow \max\{f(x^{(k)}), f(x^{(k-1)}), \ldots, f(x^{(k-K)})\}$;
  while $f(\bar{x}) \geq \hat{f} - \sigma \nabla f(x^{(k)})^T (x^{(k)} - \bar{x})$ **do**
    $\beta \leftarrow \beta^2$;
    $\bar{x} \leftarrow P_{\mathcal{Q}}(x^{(k)} - \beta\theta_k \nabla f(x^{(k)}))$;
  $x^{(k+1)} \leftarrow \bar{x}$;

B. Nesterov's optimal method

Nesterov [15] proposed a gradient-based method that for given $\mu > 0$ achieves the convergence rate

$$f(x^{(k)}) - f^\star \leq \left(1 - \sqrt{\frac{\mu}{L}}\right)^k \cdot C_{\mathrm{N}}, \tag{8}$$

where C_{N} is a constant, and he proved the method to be optimal, i.e., that no gradient-based method can achieve better worst-case convergence rate on the class of strongly convex problems.

Comparing (7) and (8), we see how the ratio μ/L affects the predicted worst-case convergence rates: When μ/L decreases, both convergence rates become slower, but less in (8) due to the square root. We therefore expect Nesterov's method to show better convergence for smaller μ/L. Small μ/L arise for instance when the number of views is small, see [8].

Nesterov's method requires that both μ and L are given by the user, and in order for the method to be convergent μ must be chosen sufficiently small and L sufficiently large. For real world applications such as CT, μ and L are seldom known, which makes the method impractical. Taking overly conservative estimates can depreciate the better convergence rate (8); hence, accurate estimates of μ and L are important.

C. Estimating μ and L

A sufficiently large L can be chosen using *back-tracking line search* [16], [17], see Algorithm 2: **BT** for pseudo-code. Essentially, an estimate $\bar{L}$ of L is increased by multiplication with a constant $\rho_L > 1$ until (5) is satisfied.

Accurately estimating μ, such that (4) is satisfied globally, is more difficult. Here, we propose a simple and computationally inexpensive heuristic: In the kth iteration choose an estimate μ_k as the largest value of μ that satisfies (4) between $x^{(k)}$ and $y^{(k)}$, and make the μ_k-sequence non-increasing:

$$\mu_k = \min \left\{ \mu_{k-1}, \frac{f(x) - f(y) - \nabla f(y)^T (x - y)}{\frac{1}{2}\|x - y\|_2^2} \right\}. \tag{9}$$

Algorithm 2: **BT**

input : $y, \bar{L}$
output: $x, \tilde{L}$
$\tilde{L} \leftarrow \bar{L}$;
$x \leftarrow P_{\mathcal{Q}}\left(y - \tilde{L}^{-1}\nabla f(y)\right)$;
while $f(x) > f(y) + \nabla f(y)^T(x-y) + \frac{1}{2}\tilde{L}\|x-y\|_2^2$ **do**
  $\tilde{L} \leftarrow \rho_L \tilde{L}$;
  $x \leftarrow P_{\mathcal{Q}}\left(y - \tilde{L}^{-1}\nabla f(y)\right)$;

Algorithm 3: **UPN**

input : $x^{(0)}, \bar{\mu}, \bar{L}$
output: $x^{(k+1)}$
$[x^{(1)}, L_0] \leftarrow \mathbf{BT}(x^{(0)}, \bar{L})$;
$\mu_0 = \bar{\mu}, \quad y^{(1)} \leftarrow x^{(1)}, \quad \theta_1 \leftarrow \sqrt{\mu_0/L_0}$;
for $k = 1, 2, \ldots$ **do**
  $[x^{(k+1)}, L_k] \leftarrow \mathbf{BT}(y^{(k)}, L_{k-1})$;
  $\mu_k \leftarrow \min\left\{\mu_{k-1}, M(x^{(k)}, y^{(k)})\right\}$;
  $\theta_{k+1} \leftarrow$ positive root of $\theta^2 = (1-\theta)\theta_k^2 + (\mu_k/L_k)\,\theta$;
  $\beta_k \leftarrow \theta_k(1-\theta_k)/(\theta_k^2 + \theta_{k+1})$;
  $y^{(k+1)} \leftarrow x^{(k+1)} + \beta_k(x^{(k+1)} - x^{(k)})$;

We call the Nesterov method equipped with estimation of μ and L Unknown Parameter Nesterov (UPN) and pseudo-code is given in Algorithm 3: **UPN**.

Unfortunately, convergence of **UPN** is not guaranteed, since the estimate (9) can be too large. However, we have found empirically that an estimate sufficient for convergence is typically effectively determined by (9).

It is possible to ensure convergence by introducing a restarting procedure [8] at the price of lowering the convergence rate bound and thereby losing optimality of the method. However, we have found empirically that the restarting procedure is seldom needed, and for realistic parameters the simple heuristic (9) is sufficient.

D. Stopping criterion

For an unconstrained convex optimization problem such as (1) the norm of the gradient is a measure of closeness to the minimizer through the first-order optimality conditions [18]. For a constrained convex optimization problem it is possible to express a similar optimality condition, namely in terms of the *gradient map* defined by

$$G_\nu(x) = \nu\left(x - P_{\mathcal{Q}}\left(x - \nu^{-1}\nabla f(x)\right)\right), \tag{10}$$

where $\nu > 0$ is a scalar. A point $x^\star$ is optimal if and only if $G_\nu(x^\star) = 0$ for any $\nu > 0$ [17]. We can use this to design a stopping criterion: Stop the algorithm after iteration k if $\|G_\nu(x^{(k)})\|_2/N \leq \epsilon$, where ϵ is a user-specified tolerance.

For an under-determined problem, e.g. in the few-view case, the objective function in (1) is nearly flat at the minimizer, which makes it difficult to determine when a sufficiently accurate reconstruction has been found. Here, the gradient map provides a simple, yet sensitive, stopping criterion.

IV. Simulation results and discussion

A. Simulation setup

At this point we emphasize that our objective is to obtain an accurate TV reconstruction while reducing the required number of gradient method iterations. We include two reconstructions merely to demonstrate that the methods indeed are successful in solving (1), thereby reconstructing the desired image. In [8] dependence of the convergence with respect to parameter variation is explored.

To demonstrate and compare the convergence of **GP**, **GPBB** and **UPN** we set up a simple test problem. As test image x_{true} we use the threedimensional FORBILD head phantom discretized into 64^3 voxels. We simulate a parallel beam geometry with view directions evenly distributed over the unit sphere. Projections are computed as the forward mapping of the discretized image subject to additive Gaussian white noise e of relative magnitude $\|e\|_2/\|Ax_{\text{true}}\|_2 = 0.01$, i.e.,

$$b = Ax_{\text{true}} + e. \tag{11}$$

We enforce nonnegativity by taking $\mathcal{Q} = \mathbb{R}_+^{64^3}$. We consider two reconstructions: A "many-view" using 55 views and a "few-view" using only 19 views of size 91^2 pixels. In the latter case A has less rows than columns, which can be shown [8] to lead to violation of the assumption on strong convexity. The iterations are continued to the tolerance $\epsilon = 10^{-8}$ is met.

B. Simulation results

Fig. 1 shows the middle (33rd) axial voxel slice through the original 3D volume together with many-view and few-view **UPN** reconstructions using $\alpha = 0.01$. Both reconstructions reproduce the orignal features accurately, except for two small features are missing in the few-view reconstruction. Fig. 2 shows the convergence of the three methods in terms of objective value $\phi^{(k)}$ relative to the true minimal objective value $\phi^\star$ as function of the iterations k. As $\phi^\star$ is unknown, we have approximated it by computing the **UPN** solution for an ϵ two orders of magnitude lower than the value used in the iterations.

In both cases we see that **UPN** converges to a satisfactory accuracy within 2000 iterations, whereas **GP** does not, and **GPBB** only does in the former case. In the many-view case **UPN** and **GPBB** both produce a significant (and comparable) acceleration over **GP**. In the few-view case, we also observe acceleration for both, but **UPN** stands out with much faster convergence. This is in accordance with the expectation stated in Section III-B.

The adequacy of the stopping criterion is evaluated by a simple visual comparison of the few-view simulation gradient map norm decay (Fig. 2 right) and the objective decay (Fig. 2 center). Apart from the erratic decay for **GPBB** (which is caused by highly irregular step length selection) there is a pronounced correspondence, and we therefore consider the stopping criterion effective.

Although **UPN** was designed for strongly convex problems, we conclude that the method also works in the non-strongly convex case of having few-view data – in fact, from the preliminary results the non-strongly convex case is where

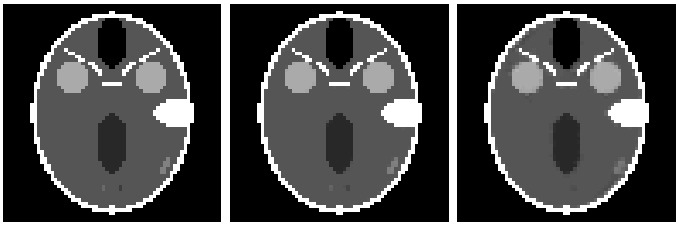


Fig. 1. Central axial slices. Left: Original. Center: Many-view **UPN** reconstruction. Right: Few-view **UPN** reconstruction. The display color range is set to $[1.04, 1.07]$ for improved viewing contrast.

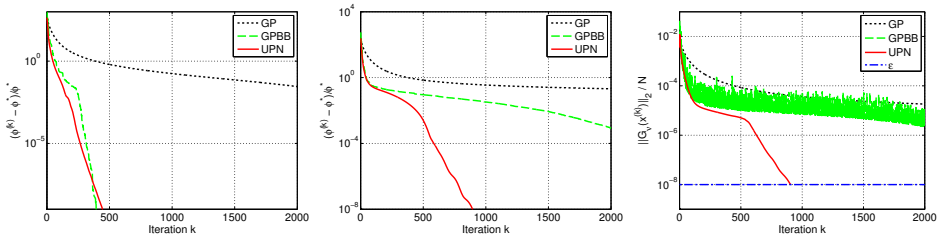


Fig. 2. Convergence histories. Left: Many-view simulation. Center: Few-view simulation. Right: Gradient map norm histories for few-view simulation.

UPN shows its biggest potential by exhibiting a much faster convergence than **GP** and **GPBB**.

V. Conclusion

We have described the gradient-based optimization methods **GPBB** and **UPN** and their worst-case convergence rates. Our simulations show that both algorithms are able to significantly accelerate high-accuracy TV-based CT image reconstruction compared to a simple gradient method. In particular **UPN** shows much faster convergence when applied to few-view data. The software (implementation in C with an interface to Matlab) is available from `http://www2.imm.dtu.dk/~pch/TVReg/`.

Acknowledgment

This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. E.Y.S. and X.P. were supported in part by NIH R01 Grant Nos. CA120540 and EB000225. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

[1] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct," *J. X-Ray Sci. Technol.*, vol. 14, pp. 119–139, 2006.

[2] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.

[3] C. R. Vogel and M. E. Oman, "Iterative methods for total variation denoising," *SIAM J. Sci. Comput.*, vol. 17, pp. 227–238, 1996.

[4] P. L. Combettes and J. Luo, "An adaptive level set method for non-differentiable constrained image recovery," *IEEE Trans. Image Proces.*, vol. 11, pp. 1295–1304, 2002.

[5] D. Goldfarb and W. Yin, "Second-order cone programming methods for total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 27, pp. 622–645, 2005.

[6] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imaging Vis.*, vol. 20, pp. 89–97, 2004.

[7] M. Hintermüller and G. Stadler, "An infeasible primal-dual algorithm for total bounded variation-based INF-convolution-type image restoration," *SIAM J. Sci. Comput.*, vol. 28, pp. 1–23, 2006.

[8] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, "Implementation of an optimal first-order method for strongly convex total variation regularization," *submitted.*

[9] A. S. Nemirovsky and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization*. Wiley-Interscience, New York, 1983.

[10] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA J. Numer. Anal.*, vol. 8, pp. 141–148, 1988.

[11] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM J. Optim.*, vol. 10, pp. 1196–1211, 2000.

[12] M. Zhu, S. J. Wright, and T. F. Chan, "Duality-based algorithms for total-variation-regularized image restoration," *Comput. Optim. Appl.*, 2008, dOI: 10.1007/s10589-008-9225-2.

[13] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for Newton's method," *SIAM J. Numer. Anal.*, vol. 23, pp. 707–716, 1986.

[14] M. Raydan, "The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem," *SIAM J. Optim.*, vol. 7, pp. 26–33, 1997.

[15] Y. Nesterov, *Introductory Lectures on Convex Optimization*. Kluwer Academic Publishers, Dordrecht, 2004.

[16] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. on Imaging Sciences*, vol. 2, pp. 183–202, 2009.

[17] L. Vandenberghe, "Optimization methods for large-scale systems," 2009, lecture Notes. www.ee.ucla.edu/ vandenbe/ee236c.html.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

Appendix L

Toward optimal X-ray flux utilization in breast CT

In Proceedings of the 11th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 359–362, Potsdam, Germany, 2011.
Proceedings available from: http://www.fully3d.org.

J. H. Jørgensen, P. C. Hansen, E. Y. Sidky, I. S. Reiser and X. Pan

Toward optimal X-ray flux utilization in breast CT

Jakob H. Jørgensen[1], Per Christian Hansen[1], Emil Y. Sidky[2], Ingrid S. Reiser[2], and Xiaochuan Pan[2]

***Abstract*—A realistic computer-simulation of a breast computed tomography (CT) system and subject is constructed. The model is used to investigate the optimal number of views for the scan given a fixed total X-ray fluence. The reconstruction algorithm is based on accurate solution to a constrained, TV-minimization problem, which has received much interest recently for sparse-view CT data.**

I. Introduction

Dose reduction has been a primary concern in diagnostic computed tomography (CT) in recent years [1]. Interest in low intensity X-ray CT is also motivated by the potential to employ CT for screening, where a large fraction of the population will be exposed to radiation dose and the majority of subjects will be asymptomatic. This abstract examines the screening application of breast CT; we simulate breast CT projection data and perform image reconstruction based on constrained, total-variation (TV) minimization. The specific question of interest is: given a fixed, total X-ray flux, what is the optimal number of views to capture in the CT scan? As the total flux is fixed, more views implies less photons per view, resulting in a higher noise level per view. On the other hand, fewer views may not provide enough sampling to recover the underlying object function. The optimal balance of these two effects will depend on the imaged subject and the imaging task. For this reason, we have focused on the breast CT application as a case study, which also has received much attention in the literature [2]–[4].

From the perspective of non-contrast CT, the breast has essentially four gray levels corresponding to: skin, fat, fibro-glandular or malignant tissue, and calcification. In designing the CT system, physical properties of the subject that are important are the complexity of the fibro-glandular tissue, which could be the limiting factor in determining the minimum number of views in the scan, and micro-calcifications and tumor spiculations, which challenge the resolution of the system.

The image reconstruction algorithm, investigated here, is based on accurate solution of constrained, TV-minimization. Constrained, TV-minimization is reconstruction by solving an optimization problem suggested in the compressive sensing (CS) community for taking advantage of sparsity of the subject's gradient magnitude [5,6]. Various algorithms based on TV-minimization have been investigated for sparse-view CT data [7]–[13], but we have also recently begun investigating TV-minimization for many-view CT with a low X-ray intensity. While the emphasis in many of these works has been algorithm efficiency, the aim here is different in that we seek accurate solution to TV-minimization in order to simplify the trade-off study. With accurate solution of TV-minimization, the resulting image can be regarded as a function of only the parameters of the optimization problem, removing the additional variability inherent in inaccurate but efficient TV-minimization solvers. The actual solver used here employs an accelerated gradient-descent algorithm which is described in an accompanying abstract and in Ref. [14,15]. This solver allows us to investigate the behavior of the solution to constrained, TV-minimization as the number of projections is varied at fixed total flux. As this is a preliminary study, the evaluation is based upon visual inspection of images obtained with a realistic computer-phantom and a CT data model incorporating physics of the low-intensity scan. Section II describes the system and subject model in detail; Sec. III briefly describes the reconstruction algorithm; and Sec. IV presents indicative results of the sampling/noise trade-off study for breast CT.

[1]Technical University of Denmark, Department of Informatics and Mathematical Modeling, Richard Petersens Plads, Building 321, 2800 Kgs. Lyngby, Denmark. [2]The University of Chicago, Department of Radiology MC-2026, 5841 S. Maryland Avenue, Chicago IL, 60637. Corresponding author: Emil Y. Sidky, E-mail: sidky@uchicago.edu

II. Breast CT model

We model the salient features of a low intensity X-ray CT system and a breast subject to gain an understanding of the trade-off between noise-per-projection and number-of-projections.

A. phantom

The breast phantom has four components: skin, fat, fibro-glandular tissue and micro-calcifications. The latter two components are the most relevant and are now described in detail. We refer all gray values to that of fat, which is taken to be 1.0. The skin gray level is set to 1.15.

Fibro-glandular tissue: The gray value is set to 1.1. The pattern of this tissue is generated by a power law noise model described in Ref. [16]. The complexity of this tissue's attenuation map is similar to what one could find in a breast CT slice. For the present study, the background fibro-glandular tissue, fat and skin are represented with as a 1024x1024 digital phantom, from which projections are computed. The reason for doing so, is that we want to isolate the issue of structural complexity of the background, while removing potential ambiguity of projection model mismatch.

Micro-calcifications: 5 small ellipses with attenuation values ranging from 1.8 to 2.1. In this case, the ellipse projections are generated from a continuous ellipse model, and unlike the rest of the phantom, these projections are not consistent with the digital projection system matrix. For these structures, object pixelization is a highly unrealistic model because of their small size; hence we employ the continuous model to generate their projection data.

The complete phantom along with a blow-up of a region of interest (ROI) containing the micro-calcifications is shown in Fig. 1. The complexity of background is apparent, and although the phantom is indeed piece-wise constant, the gradient magnitude has 55,000 non-zero values due to the structure complexity. This number is relevant for the CS argument on the accuracy of TV-minimization. While there has been no

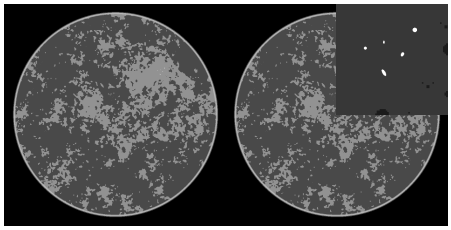


Fig. 1. Left: complete breast phantom shown in a gray scale window [0.9,1.25]. Right: same phantom with a blown-up inlay of 7.5x7.5 mm^2 ROI containing the micro-calcifications. The ROI grayscale window is [0.9,1.8]. All image reconstruction results are shown in this format.

analysis of CS recovery for CT-based system matrices, one can expect that at least twice as many samples as non-zero elements in the gradient magnitude will be needed for accurate image reconstruction with TV-minimization under noiseless conditions.

B. data model

As the primary goal of this study is to investigate a noise trade-off, the CT model includes a random component modeling the detection of finite numbers of X-ray quanta. The process of generating the simulated CT data starts with computing a noiseless sinogram:

$$g_j = \int_{L_j} d\ell\, f_{\text{digital}}[\vec{r}(\ell)] + f_{\mu\text{calc}}[\vec{r}(\ell)], \tag{1}$$

where g_j is the jth line integral of the phantom over the ray L_j with the index j running from 1 to N_{data}; N_{data} is the product of the number of projections and the number of detector bins per projection; and $f_{\text{digital}}[\vec{r}(\ell)]$ and $f_{\mu\text{calc}}[\vec{r}(\ell)]$ represent the digital and continuous components of the phantom, respectively. The measurements g_j are used for the noiseless reconstructions.

In order to include a random element to the data, which depends on N_{data} in a fairly realistic way, we compute a mean photon number per detector bin based on g_j and a total photon intensity of the scan:

$$n_j^{(\text{mean})} = \frac{N_{\text{photon}}}{N_{\text{data}}} \exp(-g_j),$$

where N_{photon} is the total number of photons in the scan and is here selected to be a value typical of mammography. Note that the model the scale factor will cause the mean number of photons per bin to decrease as the number of ray measurements increases. From $n_j^{(\text{mean})}$, a realization $\mathbf{n}_j$ is selected from a Gaussian distribution, using $n_j^{(\text{mean})}$ as the mean and variance. This Gaussian distribution closely models a Poisson distribution for large $n_j^{(\text{mean})}$. Finally, the photon number noise realization is converted back to a realization of a set of line integrals:

$$\mathbf{g}_j = -\ln\left(\frac{N_{\text{data}}}{N_{\text{photon}}}\mathbf{n}_j\right).$$

It is this data set which will be used for the noisy reconstructions below. While this model incorporates the basic idea of the noise-level trade-off, there are still limitations of the study. The incident intensity on each detector bin is assumed to be the same; no correlation with neighboring bins is considered; electronic noise in the detector is not accounted for; and reconstructions are performed from a single realization as opposed to an ensemble of realizations.

III. Image reconstruction by constrained TV-minimization

In order to perform the image reconstruction, we employ CS-motivated, constrained, TV-minimization:

$$\vec{f}^* = \text{argmin}\|\vec{f}\|_{\text{TV}} \text{ subject to } |X\vec{f} - \vec{\mathbf{g}}|^2 \leq \epsilon^2 \text{ and } \vec{f} \geq 0, \tag{2}$$

where the norm $\|\cdot\|_{\text{TV}}$ is the sum over the gradient magnitude of the image; the system matrix X represents discrete projection converting the image estimate $\vec{f}$ to a projection estimate $\vec{g}$; ϵ is a data error tolerance parameter controlling how closely the image estimate is constrained to agree with the available data; and the last constraint enforces non-negativity of the image. This optimization problem has served to aid in designing many new image reconstruction algorithms for CT. As the CT application is quite challenging, most of these algorithms do not yield the solution $\vec{f}^*(\epsilon)$ of Eq. (2), which should only depend on ϵ once the CT system parameters are fixed. As a result, these algorithms yield images which also depend on algorithm parameters. This is not necessarily a bad thing, but it becomes difficult to survey the effectiveness of Eq. (2) for various CT applications.

In applied mathematics, motivated by CS, there has been much effort in developing accurate solvers to Eq. (2), but few of these solvers can be applied to systems as large as those encountered in CT. To address this issue, we have been investigating means of accelerating gradient methods, which can be implemented for systems on the scale typical of CT. The proposed set of algorithms are described in detail in an accompanying submission to the meeting [15]. We do not discuss the algorithm here, but we point out that the optimization problem solved is modified, but equivalent to Eq. (2):

$$\vec{f}^* = \text{argmin } \alpha\|\vec{f}\|_{\text{TV}} + |X\vec{f} - \vec{\mathbf{g}}|^2 \text{ subject to } \vec{f} \geq 0, \tag{3}$$

where the data error term has been included in the objective function, leaving only positivity as a constraint. The penalty parameter α replaces the role of ϵ above. We use the accelerated gradient algorithm to solve Eq. (3) to a numerical accuracy greater than what would be visible in the images; thus, we describe the following resulting images as solutions to this optimization problem. To make the connection with the Eq. (2) is straight-forward; the corresponding ϵ to a given α is found by computing $|X\vec{f}^* - \vec{\mathbf{g}}|$ where $\vec{f}^*$ is found from Eq. (3).

IV. Results

For this initial survey of a breast CT simulation, we show two main sets of results. The first set of images are reconstructed from noiseless data for different numbers of views. The idea is to see how well TV-minimization performs in recovering the complex breast phantom under ideal conditions. The second set of images includes noise at a fixed exposure, and as described in Sec. II-B, the noise-level per projection increases with the the number of projections.

All reconstructions are performed on a 1024x1024 grid with 100 micron pixel widths. The simulated fan-beam geometry has an 80 cm source to detector distance with a circular source

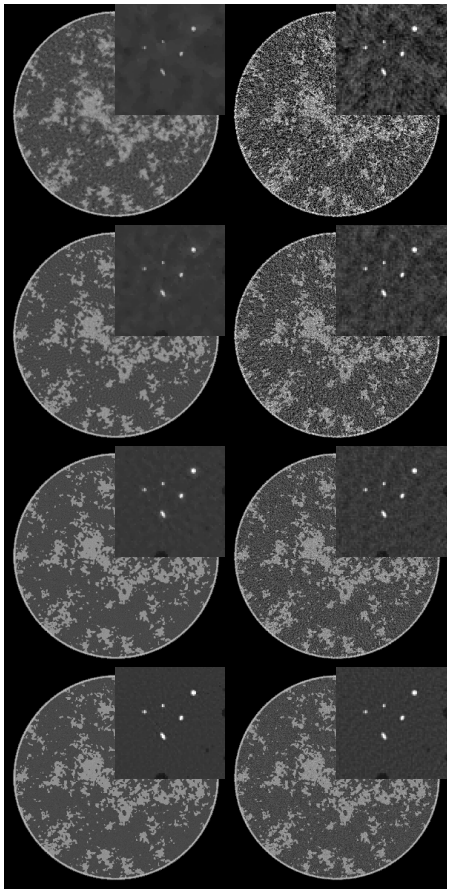


Fig. 2. Left column: images reconstructed by TV-minimization. Right column: images reconstructed by FBP. The data do not include noise, and the number of views are 64, 128, 256, and 512 going from top to bottom.

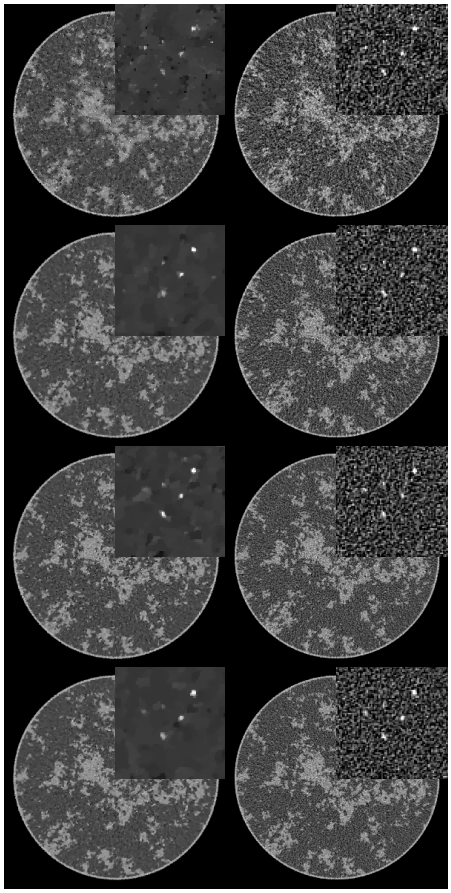


Fig. 3. Same as Fig. 2 except the noise model discussed in Sec. II-B is included.

trajectory of radius 60 cm. The detector is modeled as having 1024 detector bins, and there is no truncation in the projection data.

A. image reconstruction from noiseless data

In Fig. 2, we show images reconstructed from 64 to 512 projections for both TV-minimization and filtered back-projection (FBP). For TV-minimization in this study we set $\alpha = 10^{-6}$, which corresponds to a very tight data constraint. As noted above the sparsity of the gradient magnitude is on the order of 50,000. Accordingly, from CS-based arguments, one could only expect to start to achieve accurate reconstruction when the number of measured line integrals exceeds 100,000, which in this case means 100 projections. An important part of CS theory deals with computing the factor between the sparsity level and necessary number of measurements for accurate recovery. This factor is unknown for TV-minimization applied to the X-ray transform, but we can see from the reconstructions that the accuracy is greatly improved in going from 128 views to 256 views. There is still a perceptible improvement in the image recovery in going to 512 views, which still represents an under-determined system despite the fact that 512 views is normally not thought of as a sparse-view data set. Again, it is the complexity of the phantom which is responsible for this behavior. The accompanying FBP results give an indication on the ill-posedness of reconstruction from the various configurations with different numbers of projections.

The results for the micro-calcification ROI are interesting in that this particular feature of the image is recovered for all data sets down to the 64-projection data set. This is not too surprising because the micro-calcifications are certainly sparse

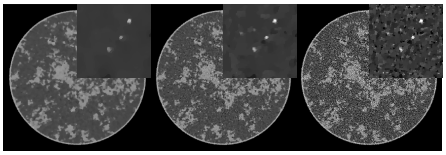


Fig. 4. Images for 512-view, noisy projection data obtained with TV-minimization for (left) $\alpha = 1 \times 10^{-3}$, (middle) $\alpha = 5 \times 10^{-4}$, and (right) $\alpha = 2 \times 10^{-4}$.

in the gradient magnitude. But this result emphasizes that the success of an image reconstruction algorithm depends also on the imaging task and the subject.

For the larger goal of determining the optimal number of views, it is clear that "structure noise" – artifacts due to the complex object function– can play a significant role for this breast phantom.

B. image reconstruction from noisy data

For the noise studies, we again investigate data sets with the view number varying between 64 and 512. For these reconstructions, α is also varied between $1. \times 10^{-6}$ and $5. \times 10^{-4}$. In Fig. 3, we show the TV-minimization images compared with FBP, as a reference. The optimal values of α for each TV-minimization image is chosen by visual inspection. The FBP fill images are smoothed by convolving with a Gaussian distribution of width 140 microns (chosen by visual inspection), and the ROI images are unregularized. While it is not too surprising that the FBP image quality appears to increase with projection number, it is somewhat surprising that the same trend is apparent for image reconstruction by TV-minimization. The 512-view data set seems to yield, visually, the optimal result in that the ROI appears to have the least amount of artifacts. While most of the micro-calcifications are visible in each reconstruction, the artifacts and noise texture in the sparse-view images can be distracting and mistaken for additional micro-calcifications. It seems that the increased noise-level per view impacts the reconstruction less than artifacts due to insufficient sampling. That we obtain this result with a CS algorithm is interesting, and warrants further investigation with more rigorous and quantitative evaluation.

To appreciate the impact of α, we focus on the 512-view data set and display images in Fig. 4 for three cases. Small α corresponds to a tight data constraint, resulting in salt-and-pepper noise in the image due to the high noise-level of the data. Increasing α reduces the image noise and eventually removes small structures.

V. CONCLUSION

We have performed a preliminary investigation of a fixed X-ray exposure trade-off between number-of-views and noise-level per view for a simulation of a breast CT system. This investigation employed a CS image reconstruction algorithm which should favor sparse-view data. Moreover, the simulated data are generated from a digital projection matched with the projector used in the image reconstruction algorithm – another factor that should favor sparse-view data. Despite this, the complexity of the subject overrides these points and it appears that the largest number of views, in the study, yields visually the optimal reconstructed images. When other physical factors are included in the data model, for example, partial volume averaging and X-ray beam polychromaticity, one can expect that this same conclusion will hold.

Extensions to the image reconstruction algorithm will address better noise modeling. One can expect an improvement in image quality by employing a weighted, quadratic data error term derived from a realistic CT noise model. As for CS-motivated image reconstruction, the breast CT system may benefit from exploiting other forms of sparsity.

VI. ACKNOWLEDGMENTS

This work is part of the project CSI: Computational Science in Imaging, supported by grant 274-07-0065 from the Danish Research Council for Technology and Production Sciences. E.Y.S. and X.P. were supported in part by NIH R01 Grant Nos. CA120540 and EB000225. I.S.R. was supported in part by NIH Grant Nos. R33 CA109963 and R21 EB8801. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] C. H. McCollough, A. N. Primak, N. Braun, J. Kofler, L. Yu, and J. Christner, "Strategies for reducing radiation dose in ct," *Radiol. Clin. N. Am.*, vol. 47, pp. 27–40, 2009.

[2] B. Chen and R. Ning, "Cone-beam volume CT breast imaging: Feasibility study," *Med. Phys.*, vol. 29, pp. 755–770, 2002.

[3] A. L. C. Kwan, J. M. Boone, K. Yang, and S. Y. Huang, "Evaluation of the spatial resolution characteristics of a cone-beam breast CT scanner," *Med. Phys.*, vol. 34, pp. 275–281, 2007.

[4] C. J. Lai, C. C. Shaw, L. Chen, M. C. Altunbas, X. Liu, T. Han, T. Wang, W. T. Yang, G. J. Whitman, and S. J. Tu, "Visibility of microcalcification in cone beam breast CT: Effects of x-ray tube voltage and radiation dose," *Med. Phys.*, vol. 34, pp. 2995–3004, 2007.

[5] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489–509, 2006.

[6] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, pp. 21–30, 2008.

[7] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J. X-ray Sci. Tech.*, vol. 14, pp. 119–139, 2006.

[8] J. Song, Q. H. Liu, G. A. Johnson, and C. T. Badea, "Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT," *Med. Phys.*, vol. 34, pp. 4476–4483, 2007.

[9] G. H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, pp. 660–663, 2008.

[10] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, pp. 4777–4807, 2008.

[11] F. Bergner, T. Berkus, M. Oelhafen, P. Kunz, T. Pan, R. Grimmer, L. Ritschl, and M. Kachelriess, "An investigation of 4D cone-beam CT algorithms for slowly rotating scanners," *Med. Phys.*, vol. 37, pp. 5044–5054, 2010.

[12] K. Choi, J. Wang, L. Zhu, T.-S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Med. Phys.*, vol. 37, pp. 5113–5125, 2010.

[13] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan, "Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT," *Phys. Med. Biol*, vol. 55, pp. 6575–6599, 2010.

[14] T. L. Jensen, J. H. Jørgensen, P. C. Hansen, and S. H. Jensen, "Implementation of an optimal first-order method for strongly convex total variation regularization," *submitted.*

[15] J. H. Jørgensen, T. L. Jensen, P. C. Hansen, S. H. Jensen, E. Y. Sidky, and X. Pan, "Accelerated gradient methods for total-variation-based ct image reconstruction," in *submitted to the 2011 International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Potsdam, Germany, 2011.

[16] I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," *Med. Phys.*, vol. 37, pp. 1591–1600, 2010.

Bibliography

[1] The Official Web Site of the Nobel Prize. http://www.nobelprize.org/. Date accessed: 26 March 2013.

[2] F. Alter, S. Durand, and J. Froment. Adapted total variation for artifact free decompression of JPEG images. *J. Math. Imaging Vision*, 23:199–211, 2005.

[3] J.-F. Aujol. Some first-order algorithms for total variation based image restoration. *J. Math. Imaging Vision*, 34:307–327, 2009.

[4] H. H. Barrett and K. J. Myers. *Foundations of Image Science*. John Wiley & Sons, Hoboken, NJ, 2004.

[5] H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers. Model observers for assessment of image quality. *Proc. Natl. Acad. Sci. USA*, 90:9758–9765, 1993.

[6] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA J. Numer. Anal.*, 8:141–148, 1988.

[7] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imag. Sci.*, 2:183–202, 2009.

[8] A. Beck and M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.*, 18:2419–2434, 2009.

[9] S. Becker, J. Bobin, and E. J. Candès. NESTA: A fast and accurate first-order method for sparse recovery. *SIAM J. Imag. Sci.*, 4:1–39, 2011.

[10] M. Beister, D. Kolditz, and W. A. Kalender. Iterative reconstruction methods in X-ray CT. *Physica Med.*, 28:94–108, 2012.

[11] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York, NY, 1982.

[12] J. Bian, J. H. Siewerdsen, X. Han, E. Y. Sidky, J. L. Prince, C. A. Pelizzari, and X. Pan. Evaluation of sparse-view reconstruction from flat-panel-detector cone-beam CT. *Phys. Med. Biol.*, 55:6575–6599, 2010.

[13] P. Blomgren and T. F. Chan. Color TV: Total variation methods for restoration of vector-valued images. *IEEE Trans. Image Process.*, 7:304–309, 1998.

[14] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, United Kingdom, 2004.

[15] K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imag. Sci.*, 3:492–526, 2010.

[16] D. J. Brenner and E. J. Hall. Computed tomography — an increasing source of radiation exposure. *N. Eng. J. Med.*, 357:2277–2284, 2007.

[17] R. A. Brooks and G. DiChiro. Statistical limitations in x-ray reconstructive tomography. *Med. Phys.*, 3:237–240, 1976.

[18] A. M. Bruckstein, D. L. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Rev.*, 51:34–81, 2009.

[19] M. H. Buonocore, W. R. Brody, and A. Macovski. A natural pixel decomposition for two-dimensional image reconstruction. *IEEE Trans. Biomed. Eng.*, 28:69–78, 1981.

[20] T. M. Buzug. *Computed Tomography. From Photon Statistics to Modern Cone-Beam CT*. Springer, Berlin, Germany, 2008.

[21] D. Calvetti and E. Somersalo. *Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*. Springer, New York, NY, 2007.

[22] E. Candès and J. Romberg. Sparsity and incoherence in compressive sampling. *Inverse Probl.*, 23:969–985, 2007.

[23] E. J. Candès and J. Romberg. Quantitative robust uncertainty principles and optimally sparse decompositions. *Found. Comput. Math.*, 6:227–254, 2006.

[24] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52:489–509, 2006.

[25] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59:1207–1223, 2006.

[26] E. J. Candes and T. Tao. Decoding by linear programming. *IEEE Trans. Inform. Theory*, 51:4203–4215, 2005.

[27] E. J. Candes and T. Tao. Near-optimal signal recovery from random projections: universal encoding strategies? *IEEE Trans. Inform. Theory*, 52:5406–5425, 2006.

[28] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, 21:21–30, 2008.

[29] A. Chambolle. An algorithm for total variation minimization and applications. *J. Math. Imaging Vision*, 20:89–97, 2004.

[30] A. Chambolle. Total variation minimization and a class of binary MRF models. In A. Rangarajan, B. Vemuri, and A. L. Yuille, editors, *Energy Minimization Methods in Computer Vision and Pattern Recognition. Lecture Notes in Computer Science*, volume 3757, pages 136–152. Springer, Berlin, 2005.

[31] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vision*, 40:120–145, 2011.

[32] T. Chan, G. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, 20:1964–1977, 1999.

[33] T. F. Chan and S. Esedoglu. Aspects of total variation regularized L1 function approximation. *SIAM J. Appl. Math.*, 65:1817–1837, 2005.

[34] T. F. Chan and J. Shen. *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. SIAM, Philadelphia, PA, 2005.

[35] R. Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Proc. Let.*, 14:707–710, 2007.

[36] G.-H. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med. Phys.*, 35:660–663, 2008.

[37] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, 20:33–61, 1998.

[38] P. L. Combettes and J. Luo. An adaptive level set method for nondifferentiable constrained image recovery. *IEEE Trans. Image Process.*, 11:1295–1304, 2002.

[39] J. Dahl, P. C. Hansen, S. H. Jensen, and T. L. Jensen. Algorithms and software for total variation image reconstruction via first-order methods. *Numer. Algorithms*, 53:67–92, 2009.

[40] J. Darbon and M. Sigelle. Image restoration with discrete constrained total variation, part I: Fast and exact optimization. *J. Math. Imaging Vision*, 26:261–276, 2006.

[41] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, LVII:1413–1457, 2004.

[42] A. B. de Gonzalez, M. Mahesh, K.-P. Kim, M. Bhargavan, R. Lewis, F. Mettler, and C. Land. Projected cancer risks from computed tomographic scans performed in the United States in 2007. *Arch. Intern. Med.*, 169:2071–2077, 2009.

[43] B. De Man and S. Basu. Distance-driven projection and backprojection in three dimensions. *Phys. Med. Biol.*, 49:2463–2475, 2004.

[44] M. Defrise and G. T. Gullberg. Image reconstruction. *Phys. Med. Biol.*, 51:R139–R154, 2006.

[45] M. Defrise, C. Vanhove, and X. Liu. An algorithm for total variation regularization in high-dimensional linear problems. *Inverse Probl.*, 27:065002, 2011.

[46] A. H. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Trans. Image Process.*, 7:204–221, 1998.

[47] D. C. Dobson and F. Santosa. Recovery of blocky images from noisy and blurred data. *SIAM J. Appl. Math.*, 56:1181–1198, 1996.

[48] D. Donoho and J. Tanner. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal processing. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 367:4273–4293, 2009.

[49] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52:1289–1306, 2006.

[50] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inform. Theory*, 47:2845–2862, 2001.

[51] J. Eckstein and D. P. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, 55:293–318, 1992.

[52] M. Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing.* Springer, New York, NY, 2010.

[53] M. Elad, P. Milanfar, and R. Rubinstein. Analysis versus synthesis in signal priors. *Inverse Probl.*, 23:947–968, 2007.

[54] I. A. Elbakri and J. A. Fessler. Statistical image reconstruction for polyenergetic X-ray computed tomography. *IEEE Trans. Med. Imaging*, 21:89–99, 2002.

[55] Y. C. Eldar and G. Kutyniok. *Compressed Sensing: Theory and Applications.* Cambridge University Press, Cambridge, United Kingdom, 2012.

[56] C. L. Epstein. *Introduction to the Mathematics of Medical Imaging.* SIAM, Philadelphia, PA, 2. edition, 2008.

[57] L. A. Feldkamp, L. C. Davis, and J. W. Kress. Practical cone-beam algorithm. *J. Opt. Soc. Amer. A*, 1:612–619, 1984.

[58] J. A. Fessler. Statistical Image Reconstruction Methods. In M. Sonka and J. M. Fitzpatrick, editors, *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*, chapter 1, pages 1–70. SPIE Press, Bellingham, WA, 2000.

[59] M. Fornasier, A. Langer, and C.-B. Schönlieb. A convergent overlapping domain decomposition method for total variation minimization. *Numer. Math.*, 116:645–685, 2010.

[60] D. Goldfarb and W. Yin. Second-order cone programming methods for total variation-based image restoration. *SIAM J. Sci. Comput.*, 27:622–645, 2005.

[61] T. Goldstein and S. Osher. The split Bregman method for L1-regularized problems. *SIAM J. Imag. Sci.*, 2:323–343, 2009.

[62] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The John Hopkins University Press, Baltimore, MD, 3. edition, 1996.

[63] R. Gordon, R. Bender, and G. T. Herman. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. *J. Theor. Biol.*, 29:471–481, 1970.

[64] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, pages 95–110. Springer, London, United Kingdom, 2008.

[65] L. Grippo, F. Lampariello, and S. Lucidi. Nonmonotone line search technique for Newton's method. *SIAM J. Numer. Anal.*, 23:707–716, 1986.

[66] J. Hadamard. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, CT, 1923.

[67] X. Han, J. Bian, D. R. Eaker, T. L. Kline, E. Y. Sidky, E. L. Ritman, and X. Pan. Algorithm-enabled low-dose micro-CT imaging. *IEEE Trans. Med. Imaging*, 30:606–620, 2011.

[68] X. Han, J. Bian, E. L. Ritman, E. Y. Sidky, and X. Pan. Optimization-based reconstruction of sparse images from few-view projections. *Phys. Med. Biol.*, 57:5245–5273, 2012.

[69] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, PA, 1998.

[70] P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, Philadelphia, PA, 2010.

[71] P. C. Hansen and M. Saxild-Hansen. AIR Tools – A MATLAB package of algebraic iterative reconstruction methods. *J. Comput. Appl. Math.*, 236:2167–2178, 2012.

[72] G. T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. Springer, London, United Kingdom, 2 edition, 2009.

[73] G. T. Herman and R. Davidi. Image reconstruction from a small number of projections. *Inverse Probl.*, 24:45011–45028, 2008.

[74] G. N. Hounsfield. Computerized transverse axial scanning (tomography): Part I. Description of system. *Brit. J. Radiol.*, 46:1016–1022, 1973.

[75] CVX Research Inc. CVX: Matlab Software for Disciplined Convex Programming, version 1.21, http://cvxr.com/cvx, 2012.

[76] S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bull. Acad. Pol. Sci. Lett.*, A35:355–357, 1937.

[77] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer, New York, NY, 2005.

[78] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging*. IEEE Press, New York, NY, 1988.

[79] W. A. Kalender. X-ray computed tomography. *Phys. Med. Biol.*, 51:R29–R43, 2006.

[80] R. M. Lewitt. Alternatives to voxels for image representation in iterative reconstruction algorithms. *Phys. Med. Biol.*, 37:705–716, 1992.

[81] M. Li, H. Yang, and H. Kudo. An accurate iterative reconstruction algorithm for sparse objects: application to 3D blood vessel reconstruction from a limited number of projections. *Phys. Med. Biol.*, 47:2599–2609, 2002.

[82] Y. Li and F. Santosa. A computational algorithm for minimizing total variation in image restoration. *IEEE Trans. Image Process.*, 5:987–995, 1996.

[83] K. K. Lindfors, J. M. Boone, M. S. Newell, and C. J. D'Orsi. Dedicated breast computed tomography: the optimal cross-sectional imaging solution? *Radiol. Clin. N. Am.*, 48:1043–1054, 2010.

[84] W. Ludwig, S. Schmidt, E. M. Lauridsen, and H. F. Poulsen. X-ray diffraction contrast tomography: a novel technique for three-dimensional grain mapping of polycrystals. I. Direct beam case. *J. Appl. Crystallogr.*, 41:302–309, 2008.

[85] MOSEK ApS. MOSEK Optimization Software, version 6.0.0.122, http://www.mosek.com,, 2011.

[86] J. L. Mueller and S. Siltanen. *Linear and Nonlinear Inverse Problems with Practical Applications*. SIAM, Philadelphia, PA, 2012.

[87] S. Nam, M. E. Davies, M. Elad, and R. Gribonval. The cosparse analysis model and algorithms. *Appl. Comput. Harmon. Anal.*, 34:30–56, 2013.

[88] F. Natterer. *The Mathematics of Computerized Tomography*. John Wiley & Sons, New York, NY, 1986.

[89] F. Natterer and F. Wübbeling. *Mathematical Methods in Image Reconstruction*. SIAM, Philadelphia, PA, 2001.

[90] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

[91] Y. Nesterov. Smooth minimization of non-smooth functions. *Math. Program.*, 103:127–152, 2005.

[92] M. Nikolova, M. K. Ng, S. Zhang, and W.-K. Ching. Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. *SIAM J. Imag. Sci.*, 1:2–25, 2008.

[93] X. Pan, E. Y. Sidky, and M. Vannier. Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction? *Inverse Probl.*, 25:123009, 2009.

[94] X. Pan, J. Siewerdsen, P. J. La Riviere, and W. A. Kalender. Anniversary paper: Development of x-ray computed tomography: The role of Medical Physics and AAPM from the 1970s to present. *Med. Phys.*, 35:3728–3739, 2008.

[95] M. Persson, D. Bone, and H. Elmqvist. Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. *Phys. Med. Biol.*, 46:853–866, 2001.

[96] M. E. Pfetsch and A. M. Tillmann. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *Arxiv preprint arXiv:1205.2081*, 2012.

[97] T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. *2011 International Conference on Computer Vision*, pages 1762–1769, 2011.

[98] H. F. Poulsen and X. Fu. Generation of grain maps by an algebraic reconstruction technique. *J. Appl. Crystallogr.*, 36:1062–1068, 2003.

[99] J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte der Sächsischen Akademie der Wissenschaft*, 69:262–277, 1917.

[100] J. Radon. On the determination of functions from their integral values along certain manifolds (translated by P. C. Parks). *IEEE Trans. Med. Imaging*, 5:170–176, 1986.

[101] S. Ramani and J. A. Fessler. A splitting-based iterative algorithm for accelerated statistical X-ray CT reconstruction. *IEEE Trans. Med. Imaging*, 31:677–688, 2012.

[102] E. A. Rashed and H. Kudo. Statistical image reconstruction from limited projection data with intensity priors. *Phys. Med. Biol.*, 57:2039–2061, 2012.

[103] M. Raydan. The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. *SIAM J. Optimiz.*, 7:26–33, 1997.

[104] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrieß. Improved total variation-based CT image reconstruction applied to clinical data. *Phys. Med. Biol.*, 56:1545–1561, 2011.

[105] L. I. Rudin, S Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60:259–268, 1992.

[106] R. L. Siddon. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med. Phys.*, 12:252–255, 1985.

[107] E. Y. Sidky, M. A. Anastasio, and X. Pan. Image reconstruction exploiting object sparsity in boundary-enhanced X-ray phase-contrast tomography. *Opt. Express*, 18:10404–10422, 2010.

[108] E. Y. Sidky, C.-M. Kao, and X. Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT. *J. X-Ray Sci. Technol.*, 14:119–139, 2006.

[109] E. Y. Sidky and X. Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Phys. Med. Biol.*, 53:4777–4807, 2008.

[110] E. Y. Sidky, X. Pan, I. S. Reiser, R. M. Nishikawa, R. H. Moore, and D. B. Kopans. Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms. *Med. Phys.*, 36:4920–4932, 2009.

[111] R. Smith-Bindman, J. Lipson, R. Marcus, K.-P. Kim, M. Mahesh, R. Gould, A. B. de Gonzalez, and D. L Miglioretti. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Arch. Intern. Med.*, 169:2078–2086, 2009.

[112] J. Song, Q. H. Liu, G. A. Johnson, and C. T. Badea. Sparseness prior based iterative image reconstruction for retrospectively gated cardiac micro-CT. *Med. Phys.*, 34:4476–4483, 2007.

[113] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Probl.*, 19:S165–S187, 2003.

[114] D. Stsepankou, A. Arns, S. K. Ng, P. Zygmanski, and J. Hesser. Evaluation of robustness of maximum likelihood cone-beam CT reconstruction with total variation regularization. *Phys. Med. Biol.*, 57:5955–5970, 2012.

[115] J. Tang, B. E. Nett, and G.-H. Chen. Performance comparison between total variation (TV)-based compressed sensing and statistical iterative reconstruction algorithms. *Phys. Med. Biol.*, 54:5781–5804, 2009.

[116] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, PA, 2005.

[117] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. V. H. Winston & Sons, Washington, D.C., 1977.

[118] J. A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Trans. Inform. Theory*, 50:2231–2242, 2004.

[119] P. Tseng. On accelerated proximal gradient methods for convex-concave optimization. *Unpublished manuscript*, 2008.

[120] E. van den Berg and M. P. Friedlander. Probing the Pareto frontier for basis pursuit solutions. *SIAM J. Sci. Comput.*, 31:890–912, 2008.

[121] L. Vandenberghe. Course notes on the subgradient method UCLA course: Optimization Methods for Large-Scale Systems. http://www.ee.ucla.edu/˜vandenbe/236C/lectures/sgmethod.pdf. Date accessed: 4 April 2013.

[122] C. R. Vogel. *Computational Methods for Inverse Problems.* SIAM, Philadelphia, PA, 2002.

[123] C. R. Vogel and M. E. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17:227–238, 1996.

[124] Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imag. Sci.*, 1:248–272, 2008.

[125] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.*, 13:600–612, 2004.

[126] P. Weiss, L. Blanc-Féraud, and G. Aubert. Efficient schemes for total variation minimization under constraints in image processing. *SIAM J. Sci. Comput.*, 31:2047–2080, 2009.

[127] M. Zhu, S. J. Wright, and T. F. Chan. Duality-based algorithms for total-variation-regularized image restoration. *Comput. Optim. Appl.*, 47:377–400, 2010.

[128] Y. Zou and X. Pan. Exact image reconstruction on PI-lines from minimum data in helical cone-beam CT. *Phys. Med. Biol.*, 49:941–595, 2004.